



INSTITUTO SUPERIOR DE ENGENHARIA DE LISBOA

**Área Departamental de Engenharia de Eletrónica
e Telecomunicações e de Computadores**



Arquitetura para Aquisição de Respostas a Perguntas Frequentes

PEDRO MIGUEL REIS TENDEIRO
(Licenciado)

Dissertação de natureza científica realizada para obtenção do grau de
Mestre em Engenharia Informática e de Computadores

Orientador:

Doutor Porfírio Pena Filipe

Júri:

Presidente: Mestre Vítor Jesus Sousa de Almeida

Vogais:

Mestre Paulo Alexandre Medeiros de Araújo
Doutor Porfírio Pena Filipe

Março de 2013

*“País desenvolvido não é onde o pobre tem carro.
É onde o rico utiliza o transporte público.”*

Resumo

Um dos objetivos da Comissão Europeia é melhorar a sustentabilidade e a acessibilidade dos transportes coletivos, promovendo a utilização de transportes mais amigos do ambiente. Neste âmbito, é importante disponibilizar informação adequada e atualizada sobre as opções de mobilidade oferecidas pelos operadores de transportes.

Devido à falta ou pouca qualidade da informação, sobre transportes públicos coletivos de passageiros, os viajantes têm dificuldade em escolher as melhores opções de mobilidade. Esta dificuldade afeta a qualidade de vida impondo adicionalmente custos ambientais e económicos.

Com o objetivo de contribuir para melhorar a qualidade da informação sobre as ofertas de mobilidade, tendencialmente mais sustentáveis, propõe-se neste trabalho que a informação prática sobre transportes, associada a entidades representadas no domínio (i.e. pontos de interesse, autoridades/operadores de transportes e eventos), assuma a forma de um conjunto de Perguntas Frequentes (**FAQ**). Para o efeito, propõe-se uma Arquitetura Orientada a Serviços (**SOA**) para suportar as FAQ e concretizar a respetiva aquisição de respostas a partir de fontes heterogéneas de informação.

Para validar experimentalmente a arquitetura proposta foi implementado um demonstrador, aplicável ao contexto da *Integra Social Network (ISN)*, que reforça a contribuição do Instituto Superior de Engenharia de Lisboa (**ISEL**) para o projeto europeu *Seamless Travel across the Atlantic area Regions using Sustainable Transport (START)*.

Abstract

One of the objectives of the European Commission is to improve the sustainability and accessibility of collective transport, while simultaneously promoting the use of more environmentally friendly transports. In this scope, it is important to make available adequate and updated information about the mobility options offered by transport operators.

Due to poor quality or lack of information, about public collective passenger transport, travellers have difficulties in choosing the best mobility options. This difficulty affects the quality of life by imposing both environmental and economic costs.

With the objective of contributing to improve the quality of information about mobility options, tendentiously more sustainable, it is proposed in this work that the practical information about transports, associated to represented domain entities (e.g. points of interest, transport authorities/operators and events), should assume the form of a set of Frequently Asked Questions (**FAQ**). To this effect, a Service Oriented Architecture (**SOA**) is proposed to model the FAQs and realize the respective acquisition of answers from heterogeneous sources of information.

To experimentally validate the proposed architecture a demonstrator was implemented, applicable to the context of Integra Social Network (**ISN**), which enhances the contribution of the High Institute of Engineering of Lisbon (**ISEL**) for the European project Seamless Travel across the Atlantic area Regions using Sustainable Transport (**START**).

Palavras-chave

Keywords

Palavras-chave

Transportes Sustentáveis, Arquitetura Orientada a Serviços, Pontos de Interesse, Rede Social, Perguntas Frequentes, Pergunta-Resposta

Keywords

Sustainable Transport, Service Oriented Architecture, Point of Interest, Social Network, Frequently Asked Question, Question Answering

Agradecimentos

A realização deste trabalho só foi possível graças à colaboração de um conjunto de pessoas, às quais agradeço, especialmente:

- Ao Professor Porfírio Filipe, pela sua orientação e por me ter proporcionado a oportunidade de trabalhar neste tema tão relevante atualmente;
- Ao Professor Paulo Martins, pelo apoio prestado no âmbito do projeto *Seamless Travel across the Atlantic area Regions using Sustainable Transport (START)*;
- Aos colegas André Costa e David Antunes, pelos esclarecimentos oportunos sobre a arquitetura da *Integra Social Network (ISN)*;
- À colega Carina Gomes pela colaboração na motivação e organização do serviço implementado;
- À minha família que nunca deixou que me faltasse nada;
- Aos meus amigos por me ajudarem a manter a sanidade mental.

Acrónimos

5W	<i>Five ‘W’ (and 1 ‘H’)</i>
DOM	<i>Data Object Model</i>
FAQ	<i>Frequently Asked Questions</i>
HTML	<i>Hypertext Markup Language</i>
ISN	<i>Integra Social Network</i>
JSON	<i>JavaScript Object Notation</i>
POI	<i>Point of Interest</i>
RDF	<i>Resource Description Framework</i>
RI	Recuperação de Informação
RS	Rede Social
SOA	<i>Service Oriented Architecture</i>
SPR	Sistema Pergunta-Resposta
START	<i>Seamless Travel across the Atlantic area Regions using Sustainable Transport</i>
XML	<i>eXtensible Markup Language</i>
WCF	<i>Windows Communication Foundation</i>
WF	<i>Windows Workflow Foundation</i>
WS	<i>Web Service(s)</i>
WSDL	<i>Web Services Description Language</i>

Convenções Tipográficas

Apresentam-se seguidamente algumas convenções tipográficas empregues na escrita deste documento:

- (i) Aplica-se no texto normal a fonte *Arial*, com tamanho 12 e espaçamento entre linhas de 1,5 cm

Exemplo: Transportes públicos;

- (ii) Aplicam-se números entre parêntesis retos para representar referências bibliográficas

Exemplo: Transportes sustentáveis [1];

- (iii) Aplica-se texto em Itálico para identificar estrangeirismos

Exemplo: *Integra Social Network*;

- (iv) Aplica-se texto em Negrito para definir acrónimos

Exemplo: *Integra Social Network* (**ISN**).

Índice

1	Introdução.....	1
1.1	Motivação.....	1
1.2	Sistemas de Informação sobre Transportes.....	2
1.3	Enquadramento.....	4
1.4	Objetivo.....	5
1.5	Organização do Documento.....	6
2	Fundamentos.....	7
2.1	Sistemas Pergunta-Resposta.....	7
2.2	Ponto de Interesse.....	9
2.3	Conceito 5W.....	11
2.4	Redes Sociais.....	13
2.4.1	História na Computação.....	13
2.4.2	Influência das RS no Quotidiano.....	14
2.5	<i>Mashups</i>	15
2.6	Arquitetura Orientada a Serviços.....	15
2.6.1	Coreografia Versus Orquestração.....	17
2.6.2	Coreografia.....	17
2.6.3	Orquestração.....	17
3	Proposta.....	19
3.1	Arquitetura para Aquisição de Respostas.....	19
3.2	Modelação de Informação Prática.....	20
3.2.1	Informação Prática.....	21

3.2.2	Palavras-Chave	23
3.2.3	Pergunta	24
3.2.4	Resposta	25
3.2.5	Fonte	28
3.2.6	Opinião do Utilizador	29
3.3	Representação de Informação Prática	30
3.3.1	Representação de Perguntas	31
3.3.2	Representação de Respostas	32
3.4	Interface de Serviços	32
4	Demonstrador	35
4.1	Análise Funcional	35
4.1.1	Casos de Utilização	35
4.1.2	Diagramas de Sequência	36
4.2	Implementação do Módulo de FAQ	37
4.3	Cenários de Teste	41
4.4	Adaptação à <i>Integra Social Network</i>	42
4.5	Simulação do Módulo na <i>Integra Social Network</i>	43
5	Epílogo	47
5.1	Balanço	47
5.2	Conclusão	47
5.3	Perspetivas de Trabalho Futuro	48
	Referências	49

Índice de Figuras

Figura 1 – Horário de Autocarro da Carris	2
Figura 2 – Horário de Comboio da CP	3
Figura 3 – Viagem Planeada pelo <i>Transporlis</i>	4
Figura 4 – Fases Comuns de um SPR	7
Figura 5 – Projeto START	8
Figura 6 – XisQuê	9
Figura 7 – Modelo de Conhecimento de POI	10
Figura 8 – Exemplo de um POI	10
Figura 9 – Conceito 5W	11
Figura 10 – Estudo Sobre Tempo Passado <i>Online</i>	14
Figura 11 – <i>Mashup</i> de Dados	15
Figura 12 – Representação de uma arquitetura SOA	16
Figura 13 – Exemplo de Coreografia	17
Figura 14 – Exemplo de Orquestração	18
Figura 15 – Arquitetura Para Aquisição de Respostas	19
Figura 16 – Modelo da Relação Entidade – Informação Prática	20
Figura 17 – Modelo do Tipo de uma Entidade	20
Figura 18 – Modelo de Informação Prática	21
Figura 19 – Modelos do Cabeçalho e Rodapé	22
Figura 20 – Modelo de uma FAQ	22
Figura 21 – Modelo de uma Pergunta	23
Figura 22 – Modelo de uma <i>Keyword</i>	24
Figura 23 – Modelo da Interrogativa	25
Figura 24 – Atributos de uma Resposta	26
Figura 25 – Classificadores de uma Resposta	26
Figura 26 – Elementos de uma Resposta	27
Figura 27 – Modelo da Fonte de Informação	28

Figura 28 – Modelo dos Argumentos	28
Figura 29 – Modelo do Sistema de Votação	29
Figura 30 – Modelo do Utilizador.....	30
Figura 31 – Interface de Serviços.....	32
Figura 32 – Casos de Utilização	35
Figura 33 – Diagrama de Sequência para a Configuração de Perguntas	36
Figura 34 – Diagrama de Sequência para a Obtenção de Respostas	37
Figura 35 - Lista de Perguntas Vazia	38
Figura 36 - Adição de uma Pergunta.....	38
Figura 37 - Adição de uma Resposta.....	39
Figura 38 - Escolha de Palavras-Chave.....	39
Figura 39 - Adição de uma Palavra-Chave	40
Figura 40 - Escolha de Perguntas.....	40
Figura 41 – Arquitetura SOA da ISN	42
Figura 42 – Configuração de Perguntas Frequentes.....	44
Figura 43 – Visualização de Perguntas Frequentes.....	45

Índice de Tabelas

Tabela 1 – Exemplo de Perguntas e Respostas	12
Tabela 2 – Reutilização de Perguntas Frequentes	12
Tabela 3 – Lista de FAQ	12
Tabela 4 – Estudo Sobre os Serviços de RS	14
Tabela 5 – Perguntas Modeladas	41
Tabela 6 – Perguntas Exemplo.....	43

Índice de Exemplos

Exemplo 1 – Relação de uma Entidade com a Informação Prática	30
Exemplo 2 – Representação de Informação Prática.....	30
Exemplo 3 – Pergunta e Respetivas Palavras-Chave	31
Exemplo 4 – Representação de uma FAQ	31
Exemplo 5 – Representação de Palavras-Chave	31
Exemplo 6 – Representação de uma Resposta.....	32

1 Introdução

Atualmente, uma das estratégias para mitigar os problemas ambientais consiste em adotar modos de transporte mais sustentáveis [1]. Com o aumento da população, aumentam também as necessidades de mobilidade. Por isso, um dos objetivos da Comissão Europeia é melhorar a sustentabilidade e a acessibilidade dos transportes públicos coletivos de passageiros, promovendo a utilização de modos de transportes amigos do ambiente [2].

Neste contexto, autoridades e operadores de transporte estão ativamente empenhados em fomentar o uso de modos de transporte sustentáveis, preferencialmente coletivos, que utilizem fontes de energia “verde”, tais como: gás natural, eletricidade ou biocombustíveis [3].

1.1 Motivação

A adoção de modos de transporte sustentáveis exige acesso, por parte dos viajantes, a fontes de informação adequada. Esta exigência agrava-se quando os viajantes desconhecem os seus destinos ou não falam a língua local.

Frequentemente, a informação disponibilizada, nomeadamente pelos operadores e autoridades de transporte, não tem a qualidade exigida pelo viajante, não contemplando aspetos, tais como: acessibilidade, atualização, autenticidade, completude, compreensibilidade, objetividade, precisão, relevância, simplicidade e valor acrescentado [4]. Adicionalmente, a informação pode ter vários formatos de apresentação em fontes que não são facilmente relacionáveis.

A qualidade da informação condiciona diretamente as decisões sobre o uso das opções de mobilidade consideradas oportunamente mais adequadas pelo viajante.

Por exemplo, alguém prepara uma viagem para um destino que desconhece, no entanto, devido a falta de informação não é possível saber antecipadamente onde comprar bilhetes, se pode transportar bagagem ou se existe localmente informação numa língua em que seja fluente.

1.2 Sistemas de Informação sobre Transportes

Os Sistemas de Informação sobre Transportes (**SIT**) tentam satisfazer as exigências de informação dos viajantes. No entanto, existem dificuldades em processar a informação oriunda de fontes heterogéneas que são tipicamente propriedade de operadores ou autoridades de transporte.

A adoção de soluções proprietárias, muitas vezes desenhadas à medida, é um obstáculo natural à interoperabilidade dificultando a obtenção de informação quando se pretende implementar processos tendencialmente automáticos.

The screenshot displays the Carris website interface for the Autocarro 208 route. At the top, there is a search bar and navigation tabs for 'Clientes' (Institucional, Media, Recrutamento, Área Pessoal) and 'Carreiras' (Simuladores de Percurso, Bilhetes e Passes, Balcão de Informações, Serviços, Contactos). The main heading is 'Autocarro 208 Cais Sodré → Estação Oriente (Interface) serviço madrugada'. Below this, there are utility buttons: 'Listar Todas', 'Ver Mapa', 'Download Ficha', 'Minhas Carreiras', and 'Calendário em vigor'. The route is identified as 'Percurso Normal'. The 'Percurso' section lists stops from Cais Sodré to Casal Vistoso. The 'Horários' section shows a table of departure times for 'Horário de Inverno'.

H	Minutos
24	30
01	00 30
02	30
03	30
04	30
05	00 35

Observações
O horário ao fim de semana é idêntico aos dias úteis.

Figura 1 – Horário de Autocarro da Carris

Habitualmente, os operadores de transporte têm o seu portal com organização e estilos próprios o que gera perplexidade nos potenciais viajantes devido à falta de homogeneidade da apresentação. São exemplo duas apresentações de informação sobre viagens, para o mesmo trajeto (Cais do Sodré – Oriente), de dois operadores, a Figura 1 relativa a autocarros do operador Companhia de Carris de Ferro de Lisboa (**Carris**) [5] e a Figura 2 relativa a comboios do operador Comboios de Portugal (**CP**) [7].

CP Passageiros A CP CP Kids

CP.pt, em linha consigo COMBOIOS DE PORTUGAL

Pesquisar

Urbanos
Viajar em Portugal
Internacional
Programas Turísticos
Viagens de Grupo
Provedor do Cliente
Cool_train

▼ Menu CP
Horários e Preços
Bilheteira
Vantagens e Promoções
Passatempos
Informação Prática
Avisos
Notícias
Newsletter
Transportes Complementares
Cultura e Lazer
Sugestões e Reclamações
myCP
▼ CP Empresas
Pedido de Adesão
Área Empresa

Horários Solicitados

Cais Do Sodre > Oriente
Entre as 00h00 e as 24h00
30 de setembro de 2012

Foram encontrados 12 resultados

Imprimir

Serviços:
U - Urbanos
AP - Alfa Pendular
IC - Intercidades
R - Regional
IR - Inter-Regional
IN - Internacional
TC - Transporte Complementar

A pesquisa de horários apresenta sempre as ligações mais rápidas para o percurso pretendido, conforme a opção de serviço selecionada. As ligações entre comboios só são asseguradas em condições normais de circulação, não sendo apresentados nesta pesquisa os comboios com ligações inferiores a 5 minutos (exceção dos Comboios Urbanos de Lisboa de 3 min). Consulte sempre o detalhe da viagem. Os preços não dispensam a sua confirmação nos postos de venda da CP.

Avisos Greve - 1 a 4 de outubro de 2012

Resultados da Ida

Serviço	Partida	Chegada	Duração	Preço	Comprar
1. TC R detalhe	Cais Do Sodre 7h00	Oriente 7h55	0h55	1 ^a /2 ^a *	ver
TC	7h00 Cais Do Sodre	7h30 Santa Apolonia	0h30	€NA	
R	7h48 Santa Apolonia	7h55 Oriente	0h07	€1,40	
2. TC R detalhe	9h00	9h55	0h55	ver	
3. TC R detalhe	12h00	12h55	0h55	ver	
4. TC R detalhe	14h00	14h55	0h55	ver	
5. TC R detalhe	15h40	16h23	0h43	ver	
6. TC R detalhe	16h00	16h55	0h55	ver	
7. TC R detalhe	17h00	17h55	0h55	ver	
8. TC R detalhe	18h00	18h55	0h55	ver	
9. TC R detalhe	19h00	19h55	0h55	ver	
10. TC R detalhe	20h00	20h55	0h55	ver	
11. TC R detalhe	22h00	22h55	0h55	ver	
12. TC R detalhe	23h30	*00h22	0h52	ver	

Legenda:
* - Hora de chegada no dia seguinte ao definido na pesquisa.
€A - O valor deste trajeto/percurso está incluído no preço anterior.
€P - O valor deste trajeto/percurso está incluído no preço seguinte.
€NA - Não Aplicável.
€ND - Preço não disponível no momento.

Info & Vendas
808 208 208
Chamadas Internacionais
(+351) 707 201 280

Nova pesquisa

Figura 2 – Horário de Comboio da CP

Mesmo assim, as dificuldades têm vindo a ser ultrapassadas, concebendo sistemas de informação cada vez mais sofisticados que aplicam abordagens com inteligência artificial, nomeadamente: planeamento automático ou semiautomático de viagens [8][9][10][11].

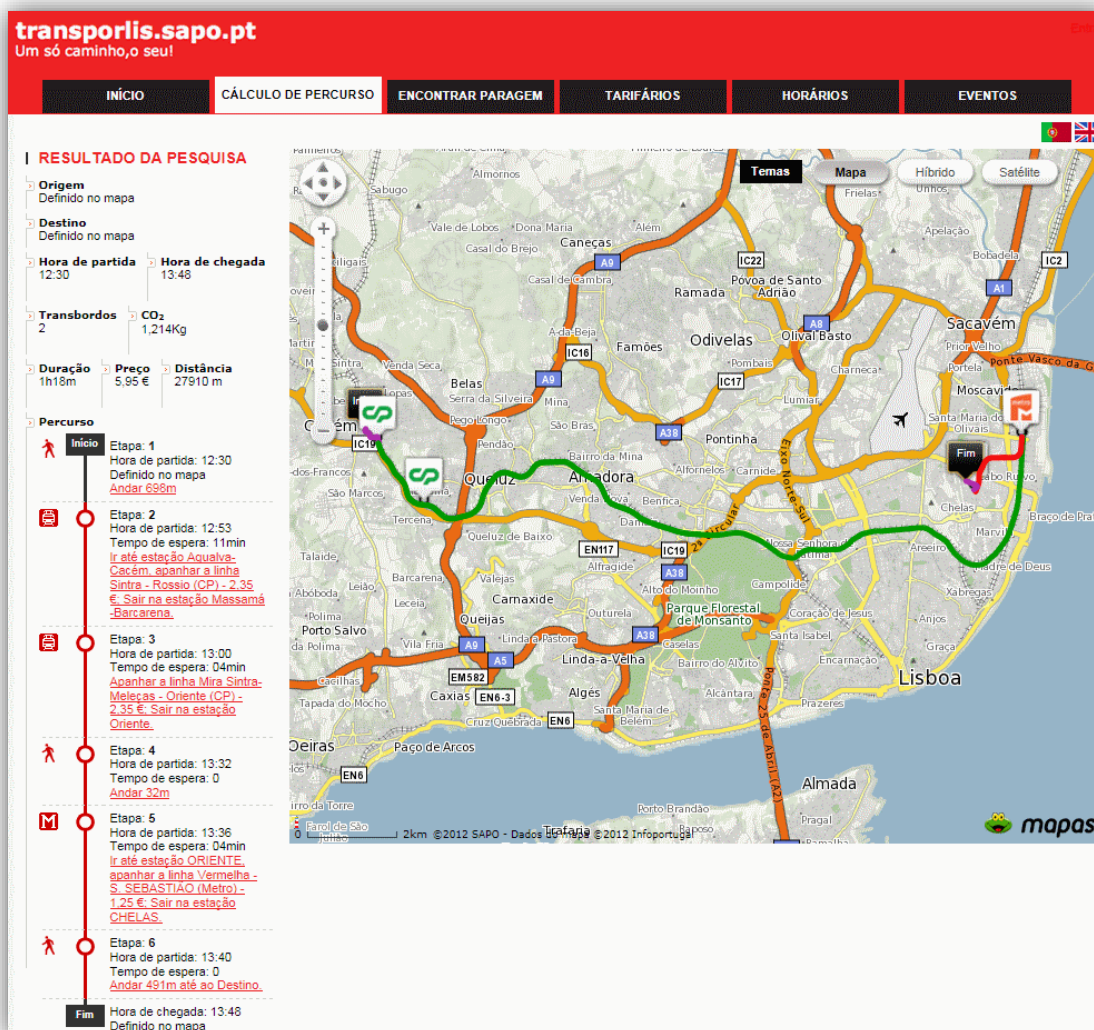


Figura 3 – Viagem Planeada pelo Transporlis

Alguns destes sistemas têm interfaces simples indicando, por exemplo, apenas a localização de paragens relativas a um modo de transporte, enquanto outros disponibilizam informação integrada relativa a múltiplos modos e operadores [12], como o apresentado na Figura 3 onde é apresentada informação relativa ao planeamento de uma viagem multimodal, do Cacém até Chelas, com dois transbordos e com duração de 1 hora e 18 minutos.

1.3 Enquadramento

Considerando a oferta dos atuais SIT identificam-se lacunas na informação que é considerada oportunamente relevante para os viajantes, principalmente para aqueles que viajam para um destino que desconhecem. O viajante necessita

frequentemente de saber, por exemplo, como: comprar e validar o título de transporte, reconhecer o meio de transporte, entrar e sair do meio de transporte, assinalar ao condutor de um autocarro que deve parar ou qual a língua da informação usada a bordo (ou nas paragens). A este tipo de informação que foca, oportunamente, aspetos práticos comuns, atribuiu-se no âmbito deste trabalho a designação de informação prática.

A informação prática está naturalmente associada a entidades existentes no domínio com diferentes granularidades, por exemplo: pontos de interesse (i.e. museu, aeroporto, vila ou cidade), autoridades/operadores de transporte ou eventos. Neste âmbito, as entidades estão relacionadas entre si devido, por exemplo, à proximidade ou à categoria.

Sendo assim, a contribuição deste trabalho foca a modelação de informação prática agregada a entidades de um domínio. Adicionalmente propõe-se que a informação prática seja modelada sob a forma de um conjunto de perguntas frequentes.

1.4 Objetivo

O objetivo deste trabalho é propor uma arquitetura computacional para suportar a aquisição de respostas a perguntas frequentes num determinado domínio, a partir de fontes heterogéneas de informação. Neste âmbito, propõe-se um modelo para representar a informação prática associada às entidades representadas num domínio. Mais concretamente trata-se o domínio da mobilidade sustentável destacando aspetos relacionados com a facilidade de acesso a transportes coletivos de passageiros.

Para validar experimentalmente a arquitetura proposta implementou-se um demonstrador, aplicável à *Integra Social Network (ISN)* [13], que reforça a contribuição do Instituto Superior de Engenharia de Lisboa (**ISEL**) [14] para o projeto europeu *Seamless Travel across the Atlantic area Regions using Sustainable Transport (START)* [15].

1.5 Organização do Documento

Este documento está organizado em 5 capítulos com a seguinte disposição:

No primeiro capítulo é realizada uma introdução à problemática da dificuldade em encontrar informação sobre transportes. Descreve-se a motivação, os SIT existentes, o enquadramento na problemática e o objetivo do trabalho. No final é apresentada a organização deste documento.

No segundo capítulo é apresentado um estudo sobre os fundamentos do trabalho, nomeadamente os Sistemas Pergunta-Resposta (**SPR**), a definição de Ponto de Interesse (**POI**), o Conceito 5W, as Redes Sociais (**RS**), *Mashups* e Arquiteturas Orientadas a Serviços (**SOA**).

No terceiro capítulo é descrita a proposta da arquitetura para disponibilizar Informação Prática e a sua possível adaptação à *Integra Social Network (ISN)*, assim como a respetiva análise funcional.

No quarto capítulo é explicado o ambiente experimental focando o demonstrador no âmbito dos serviços disponibilizados pela ISN.

No quinto capítulo são tecidas considerações finais, apresentando também perspetivas de desenvolvimento futuro.

nomeadamente: Análise da Pergunta (onde são extraídos conceitos), a Recolha de Passagens (onde os conceitos são utilizados para escolher potenciais passagens onde a resposta poderá estar) e finalmente a Extração da Resposta (a ser mostrada ao utilizador).

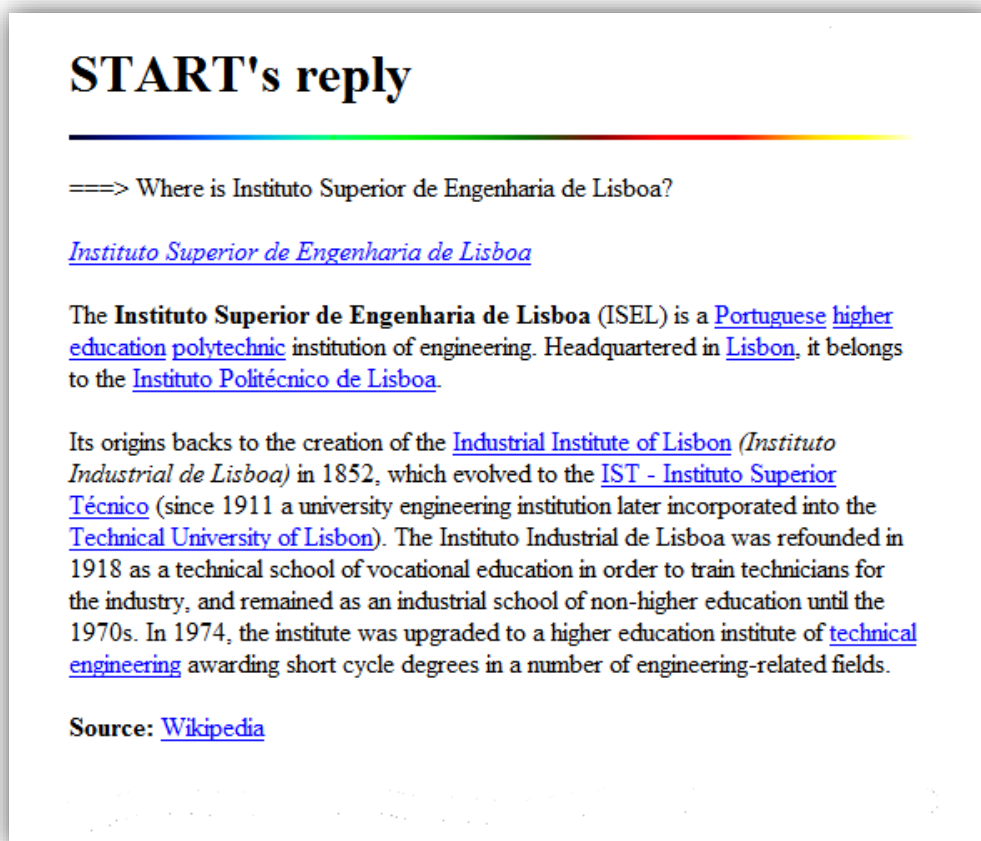


Figura 5 – Projeto START

A título de exemplo, referem-se o projeto START [20], desenvolvido no *Massachusetts Institute of Technology* (MIT), descrito na Figura 5, e o *XisQuê* [21], desenvolvido no Departamento de Informática da Universidade de Lisboa, cuja página de entrada surge na Figura 6.

Estes sistemas apresentam diferentes abordagens na obtenção da resposta. Para efeitos de demonstração do funcionamento foi indicada a mesma pergunta, em inglês e em português respetivamente, “Onde fica o Instituto Superior de Engenharia de Lisboa?”.



XisQuê Introduza uma pergunta:

Resposta a perguntas na Web portuguesa

Perguntar Limpar

Pergunta: Onde fica o Instituto Superior de Engenharia de Lisboa?

Resposta #1: Portugal
O Instituto Superior de Engenharia de Lisboa é uma das mais antigas escolas de engenharia em **Portugal**. Teve a sua génese em 30 de Dezembro de 1852, por Decreto Régio de D. Maria II, que criou o Instituto Industrial de Lisboa. Este Instituto passou a ter estatuto de ensino superior com a denominação de Instituto Superior de Engenharia de Lisboa, por Decreto-Lei 830/74 de 31 de Dezembro.
[documento fonte](#)

Resposta #2: Lisboa
O Instituto Superior de Engenharia de Lisboa (ISEL) é uma escola de engenharia em Lisboa, Portugal.
[documento fonte](#)

Figura 6 – XisQuê

O projeto START retornou um excerto de um artigo oriundo da *Wikipedia* sobre o ISEL, onde menciona a localização. Já o *XisQuê* retornou várias alternativas, sendo que as apresentadas indicam como respostas mais prováveis Portugal e Lisboa.

2.2 Ponto de Interesse

Para a resolução do problema que se levanta com a inexistência/insuficiência de informação prática, foram desenvolvidas especificações técnicas nas quais todos os locais ou eventos que suscitam interesse do público são designados por *Points of Interest (POI)* e têm um subconjunto de atributos que lhes são inerentes [22].

Por exemplo, no modelo da Figura 7 são assinalados dados relevantes acerca de um POI. Esses dados incluem um identificador, morada, coordenadas GPS, contactos, horários de funcionamento e até outros POI.

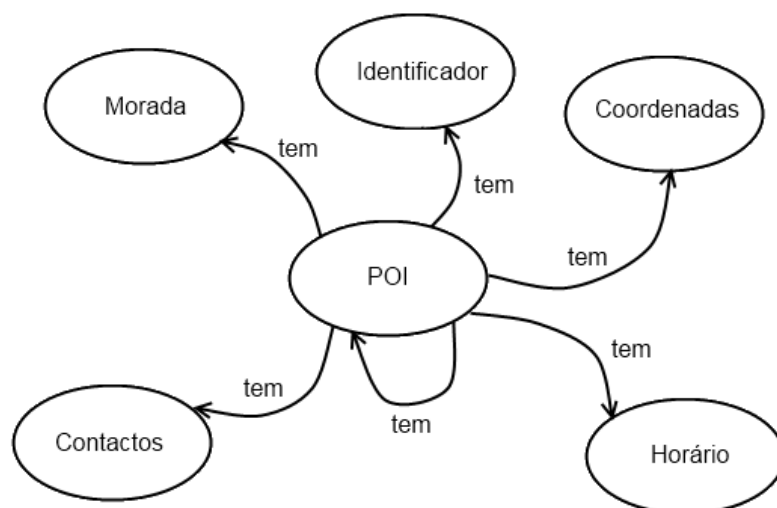


Figura 7 – Modelo de Conhecimento de POI

Este tipo de abordagem permite a existência de um POI com vários níveis de granularidade.

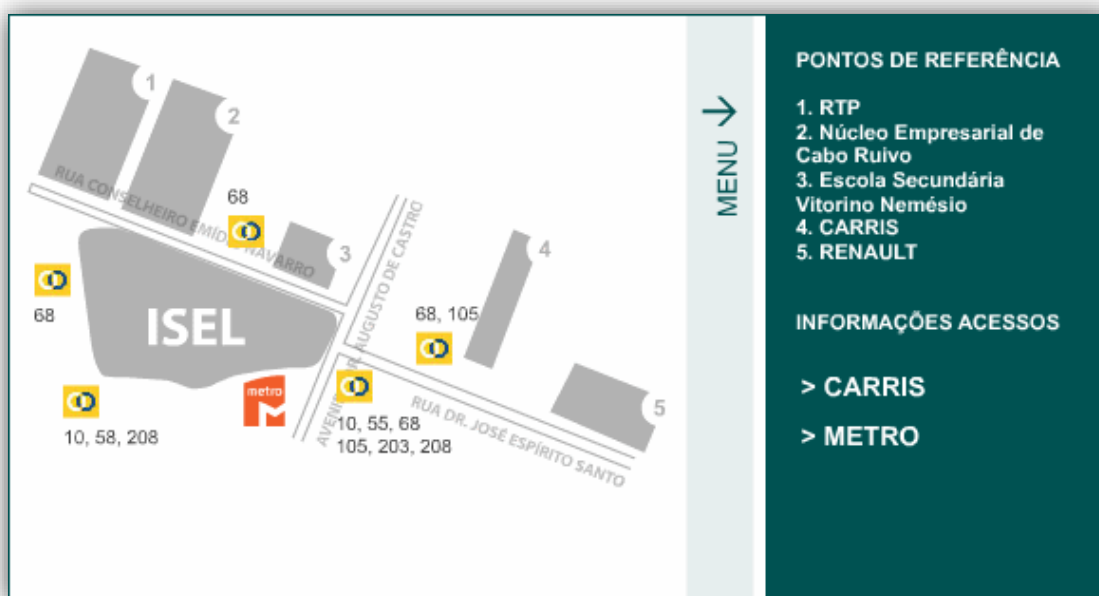


Figura 8 – Exemplo de um POI

Por exemplo, na Figura 8 apresenta-se o mapa da área envolvente do ISEL, como sendo um POI que contém um subconjunto de POI, neste caso várias paragens de autocarros da Carris [5], a estação de Metro [6], bem como vários pontos de referência.

2.3 Conceito 5W

Consultar a informação sobre um POI não é muito acessível. Por exemplo, alguém que não conheça a zona ou não possua um sistema de navegação com GPS, não pode fazer grande uso da informação geográfica. Por isso, existe interesse em fatorizar a informação em perguntas simples e reutilizáveis, cuja resposta depende do contexto onde a pergunta ao qual a pergunta está associada. Para este efeito, refere-se a importância do conceito **5W**, ilustrado na Figura 9, que corresponde em inglês aos pronomes interrogativos *Who* - Quem, *Where* - Onde, *When* - Quando, *What* - Quê, *Why* - Porquê e *How* - Como. No âmbito deste trabalho, o pronome 'Porquê' não é utilizável, porque está afeto ao utilizador, ou seja, é da responsabilidade do utilizador responder o que nunca acontece. O conceito 5W é usado por jornalistas e investigadores para tentarem responder aos assuntos em estudo.



Figura 9 – Conceito 5W

Com a utilização do 5W pretende-se transformar a informação prática sobre transportes de um determinado local ou evento em perguntas frequentes, aumentando assim a usabilidade uma vez que é oferecido um modelo de interação bem conhecido do utilizador. Estas perguntas irão permitir vários níveis de refinamento, a ponto da mesma pergunta poder ser usada para um local tão grande como um país, sendo refinada para uma cidade, ponto de interesse na cidade ou até um evento. Em relação a respostas, também é possível definir repostas heterogéneas, através do uso de serviços ou ligações para locais onde

estará efetivamente a resposta. Na Tabela 1 são exemplificadas três perguntas com os tipos de resposta possíveis.

#	Pergunta frequente	Tipo de resposta
1	Quais as coordenadas GPS?	Texto simples
2	Como chegar do centro da cidade?	Ligação a um planeador de viagens
3	Como chegar por Metro/Comboio/Autocarro?	Indicação da estação e imagem da rede de transportes respetiva

Tabela 1 – Exemplo de Perguntas e Respostas

Através da primeira e terceira perguntas são exemplificadas respostas estáticas. Na segunda pergunta recorre-se ao uso de conteúdo dinâmico.

Pergunta	Estação de comboio	Paragem de autocarro
Onde comprar bilhete?	Na bilheteira	A bordo
Quando validar bilhete?	Na plataforma	A bordo
Como pagar?	Com MB ou Dinheiro	Dinheiro

Tabela 2 – Reutilização de Perguntas Frequentes

A reutilização das perguntas para gerar respostas diferentes conforme o contexto onde são usadas, indicam-se exemplos na Tabela 2.

Pergunta	Resposta
É possível o transporte de bicicletas?	“sim / não”
É possível adquirir bilhete a bordo?	“sim / não”
Qual a frequência do transporte?	“valor”
Qual o ponto de partida?	“coordenadas, nome do local”
Qual o horário de partida?	“hora”
É necessária quantia exata?	“sim / não”
Como entrar e sair do veículo?	“frase”
Qual a linguagem a bordo?	“pt, es, fr, uk”
Qual é o portal da operadora?	“link”
Existe espaço para bagagem a bordo?	“sim / não”
Qual o custo mínimo do bilhete?	“valor” “moeda”
Existe informação em tempo real a bordo?	“sim / não”
Portal do posto de turismo?	“link”
Onde e como entrar no veículo?	“frase”

Tabela 3 – Lista de FAQ

Para manter a clareza e simplicidade, embora as perguntas permitam respostas verbosas, é boa prática mantê-las concisas. Sempre que possível deve optar-se por perguntas com respostas binárias “Sim/Não”. Uma lista de exemplos de Perguntas Frequentes (**FAQ**) sobre transportes públicos é apresentada na Tabela 3.

2.4 Redes Sociais

Uma Rede Social (**RS**), do ponto de vista de sociologia, é uma estrutura composta por atores, sendo estes pessoas ou organizações, que partilham valores e objetivos comuns. Uma das características principais é a possibilidade de relações horizontais e não hierárquicas entre os participantes [23].

Na computação, um serviço de RS é uma plataforma, ou página que facilita a construção de redes ou relações sociais entre pessoas que, por exemplo, partilham interesses, atividades, experiências ou ligações na vida real. Este tipo de serviço tem atraído milhões de pessoas, as quais os integraram no seu dia-a-dia [24].

2.4.1 História na Computação

O primeiro serviço de RS identificável como tal começou em 1997, com a rede SixDegrees.com. Tal como a maioria das RS, este permitia em simultâneo a criação de um perfil público ou semipúblico e a identificação dos amigos na rede, ao contrário de outros serviços mais antigos que só permitiam uma das duas.

Após o serviço *SixDegrees.com*, vários serviços foram aparecendo ao longo dos anos, sendo dos mais conhecidos em Portugal o *hi5* [25], *LinkedIn* [26], *Orkut* [27], *Twitter* [28], *Myspace* [29], *Facebook* [30] e mais recentemente o *Google+* [31]. Estes dois últimos serviços permitem um alto nível de modularização, na medida em que é possível adicionar aplicações na própria RS, ou associar a RS a outras aplicações. Esta última permite a criação de contas noutros serviços, sejam eles RS ou não, com a partilha dos dados pessoais dos utilizadores, através do aval positivo destes.

2.4.2 Influência das RS no Quotidiano

Em novembro de 2011, um estudo levado a cabo pela empresa *ComScore* [32], chegou aos resultados apresentados na Tabela 4.

Serviço	Visitantes Únicos (milhares)	Porcentagem
Facebook.com	792,999	55.1 %
Twitter.com	167,903	11.7 %
LinkedIn.com	94,823	6.6 %
Google+	66,756	4.6 %
MySpace	61,037	4.2 %
Outros	255,539	17.8 %
Total	1,438,877	100 %

Tabela 4 – Estudo Sobre os Serviços de RS

Um estudo mais recente, também efetuado pela mesma entidade [33] sobre o tempo passado *online* produziu o gráfico apresentado na Figura 10.

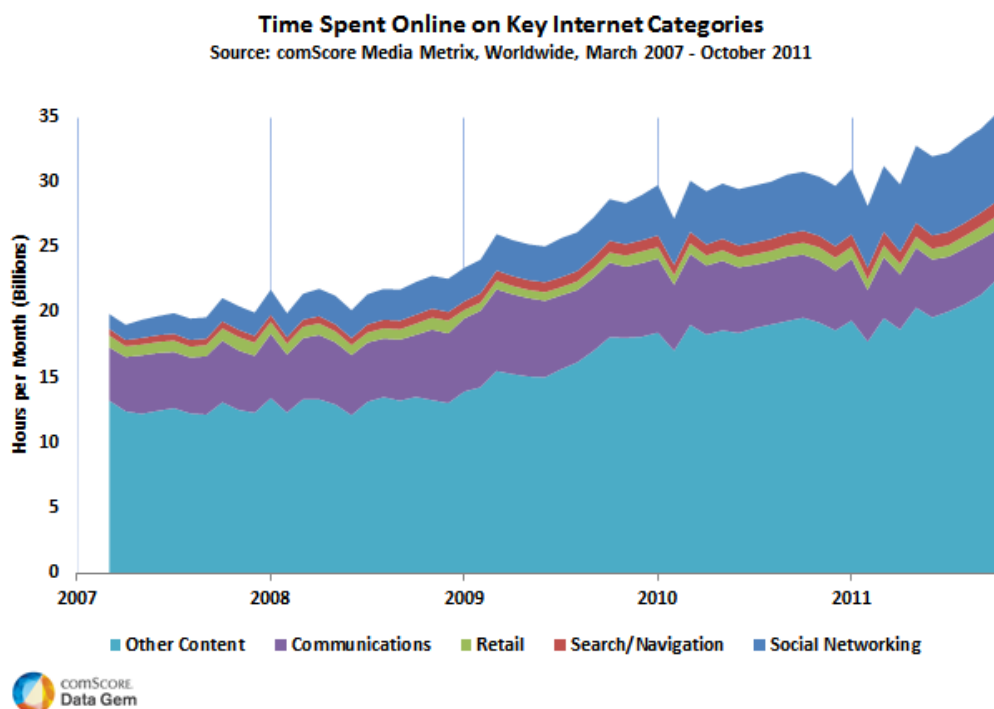


Figura 10 – Estudo Sobre Tempo Passado *Online*

De acordo com o gráfico da Figura 10, o tempo acumulado passado *online* nas várias categorias quase que duplicou entre 2007 e 2011. No entanto, foi na gama das RS que houve um aumento mais significativo.

2.5 Mashups

Um *Mashup*, embora não tenha uma definição formal [34], é no fundo uma aplicação ou página *web* que recolhe, combina e apresenta informação oriunda de fontes heterogêneas e esparsas, gerando algo novo. A Figura 11 ilustra um exemplo do funcionamento típico de um *Mashup* visto numa perspectiva *bottom-up*.

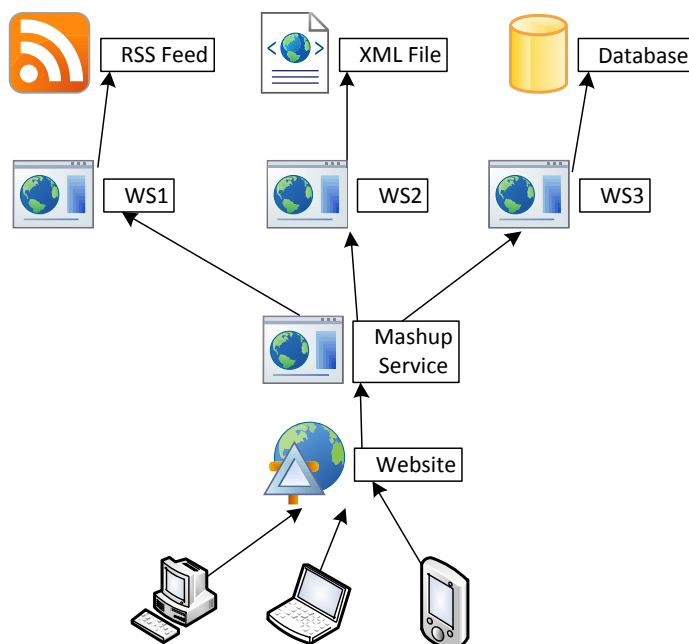


Figura 11 – *Mashup* de Dados

Ao aceder a uma página *web*, é feito um pedido ao serviço principal, na figura identificado como *Mashup Service*. Este serviço acede a um número de serviços específicos que recolhem a informação das diferentes fontes de dados [35]. Uma vez recolhida a informação, esta é retornada ao *Mashup Service* que agrega toda a informação que retorna à fonte do pedido.

2.6 Arquitetura Orientada a Serviços

Nos últimos 25 anos, as arquiteturas de *software* têm-se afirmado como uma área de saber importante para a evolução humana. Recentemente, o desacoplamento das interfaces dos processos aumentou devido ao uso de *standards* abertos que permitem desacoplar a interface de um serviço da sua

implementação. Este estilo de arquitetura veio a ser conhecido por *Service Oriented Architecture - Arquitetura Orientada a Serviços (SOA)* [36].

A implementação de SOA mais usada recorre a *Web Services (WS)* [37], normalmente definidos como sendo *software* desenhado para suportar interoperabilidade entre máquinas numa rede. Os serviços disponibilizados comunicam através de *standards* comuns na Internet, tais como o *eXtensible Markup Language (XML)*, *Web Services Description Language (WSDL)*, *Resource Description Framework (RDF)* ou *JavaScript Object Notation (JSON)*.

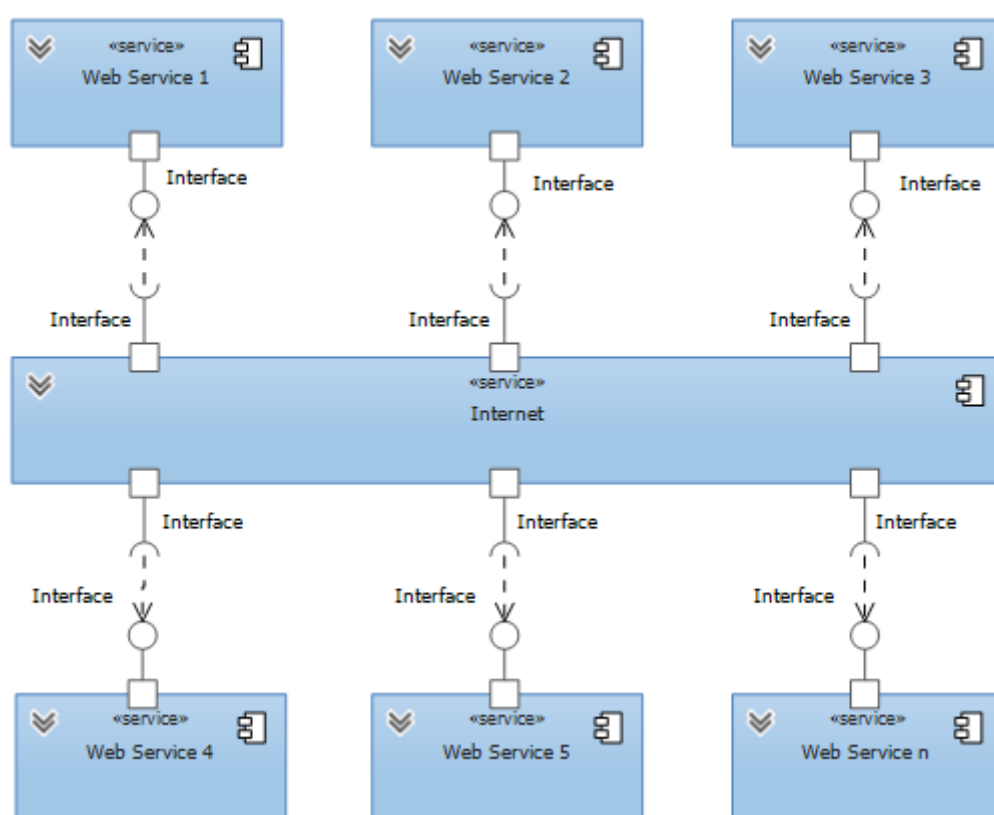


Figura 12 – Representação de uma arquitetura SOA

A integração de fontes de informação heterogêneas requer uma estratégia tecnologicamente independente para recolher os dados, independentemente da sua estrutura. Cada fonte de informação tem normalmente atribuído um WS cujo objetivo é apenas aceder, extrair e normalizar conteúdos específicos de uma fonte de informação, possibilitando assim a junção de dados normalmente incompatíveis. Ao moldar um WS a uma interface conhecida, é possível aumentar o número de fontes de dados sem que isso envolva adaptação por

parte das aplicações cliente. A representação típica deste tipo de arquitetura está surge na Figura 12, onde existe um *BUS* de serviços suportado na *Internet*.

2.6.1 Coreografia Versus Orquestração

Tendo em adoção de uma arquitetura orientada a serviços, destacam-se as combinações mais comuns de WS, designadamente: coreografia e orquestração.

Uma coreografia e uma orquestração [38] são distinguíveis pela medida em que uma coreografia é um sistema colaborativo entre WS e uma orquestração é um sistema de execução de processos. O facto de uma coreografia ser um sistema colaborativo implica que todos os WS estão interligados e colaboram entre si. Já na orquestração, tem um coordenador central que é responsável pela ordem de execução dos WS, o que invalida a necessidade de ligação entre eles.

2.6.2 Coreografia

A organização de serviços em coreografia é ilustrada na Figura 13.

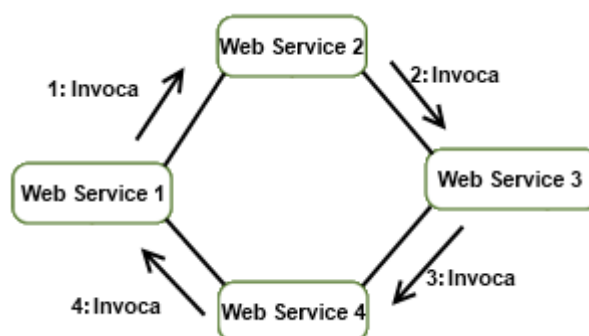


Figura 13 – Exemplo de Coreografia

Esta organização põe em risco a reutilização dos serviços por outros, porque o comportamento está pré-definido, ou seja, caso se pretenda, por exemplo, usar apenas um dos serviços não é possível devido a este invocar sempre outro serviço, ou seja, por ser um circuito fechado.

2.6.3 Orquestração

À semelhança da coreografia, é apresentada na Figura 14.uma organização de serviços sob forma de orquestração.

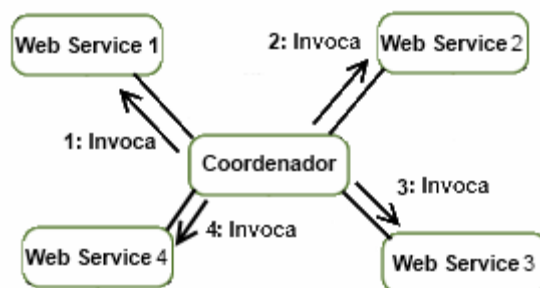


Figura 14 – Exemplo de Orquestração

Neste exemplo, é o coordenador que faz a composição dos WS, o que facilita tanto a expansão do sistema e a reutilização dos serviços, pois estes são independentes entre si.

3 Proposta

Neste capítulo é descrita a arquitetura proposta para a aquisição de respostas assim como o modelo que suporta a definição de informação prática sobre um domínio. A informação prática é acedida por um conjunto de perguntas frequentes sendo constituída pelo conjunto das respetivas respostas.

3.1 Arquitetura para Aquisição de Respostas

Devido á quantidade de serviços envolvidos e às suas diferentes naturezas foi adotada a técnica *Mashup* sobre SOA para definir a arquitetura para a aquisição de respostas.

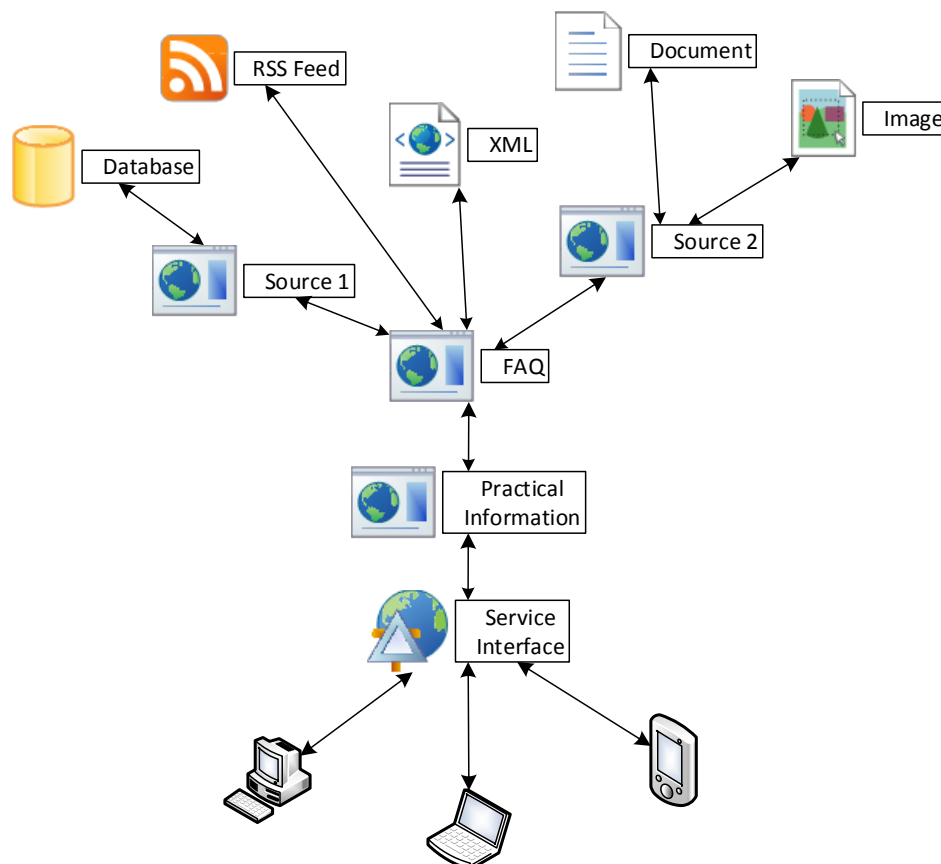


Figura 15 – Arquitetura Para Aquisição de Respostas

A combinação de *Mashup* com SOA permite maior flexibilidade para recolher informação de fontes heterogêneas facilitando o desacoplamento e a extensibilidade. Na Figura 15 ilustra-se a arquitetura para aquisição de respostas, onde o módulo de FAQ faz a orquestração dos vários serviços para adquirir informação que é posteriormente agregada e enviada para o cliente.

3.2 Modelação de Informação Prática

A modelação de informação prática sobre o domínio está associada, tal como é ilustrado na Figura 16, às entidades por uma relação.

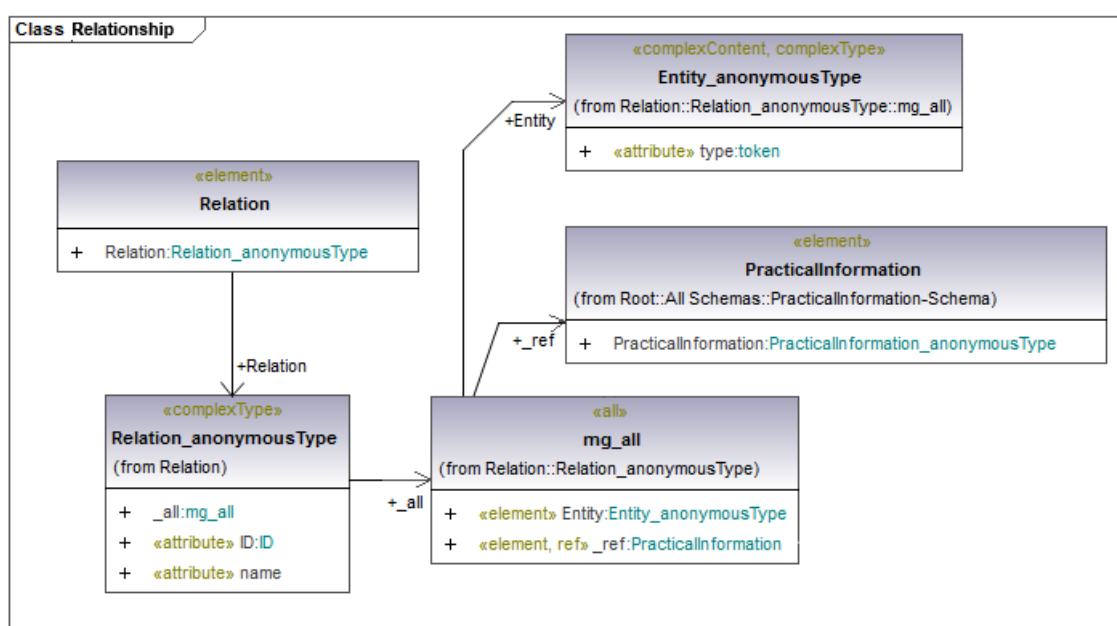


Figura 16 – Modelo da Relação Entidade – Informação Prática

A relação entre uma entidade “*Entity*” e a respetiva informação prática, “*PracticalInformation*”, tem dois atributos: um identificador único no domínio “*ID*” e um nome legível “*name*”.

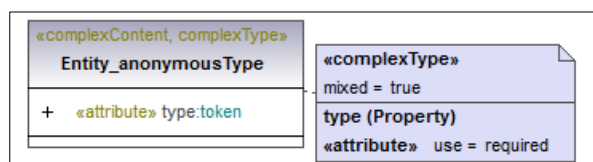


Figura 17 – Modelo do Tipo de uma Entidade

A entidade do domínio, representada na Figura 17, toma a forma de um elemento previamente conhecido, de um determinado tipo “*type*”, que o categoriza perante o sistema que posteriormente saberá como o apresentar.

3.2.1 Informação Prática

A Figura 18 esboça o modelo proposto para representar a informação prática associada a uma entidade existente no domínio.

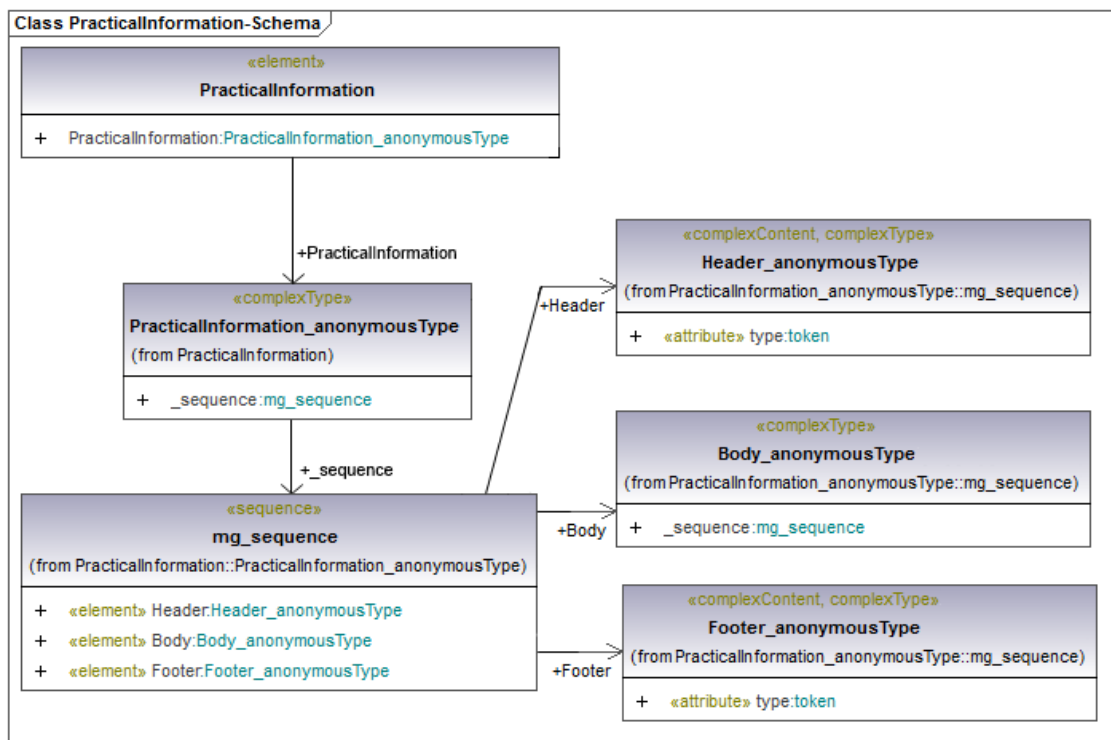


Figura 18 – Modelo de Informação Prática

A informação prática é organizada por uma sequência de três elementos, o cabeçalho “*Header*”, o corpo “*Body*” e o rodapé “*Footer*”. O cabeçalho e o rodapé destacam-se do corpo da informação prática, delimitando-a. Esta delimitação permite uma contextualização da informação através de elementos simples.

3.2.1.1 Cabeçalho e Rodapé

O cabeçalho e o rodapé, expostos na Figura 19, são elementos multimédia identificados através do seu tipo “*type*”.

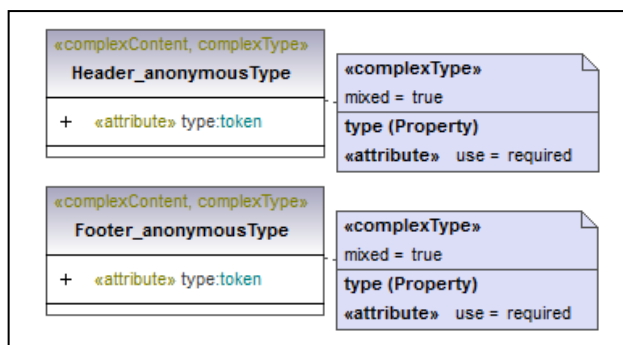


Figura 19 – Modelos do Cabeçalho e Rodapé

O atributo “*type*” dá a indicação do tipo do elemento incluído, por exemplo, se é um texto ou uma imagem, possibilitando assim modelar os conteúdos associados ao tipo de elemento.

3.2.1.2 Corpo

O corpo da informação prática, tal como ilustrado na Figura 20, é constituído por um elemento “FAQ” que por sua vez é constituído por uma sequência de elementos “Question” que representam perguntas.

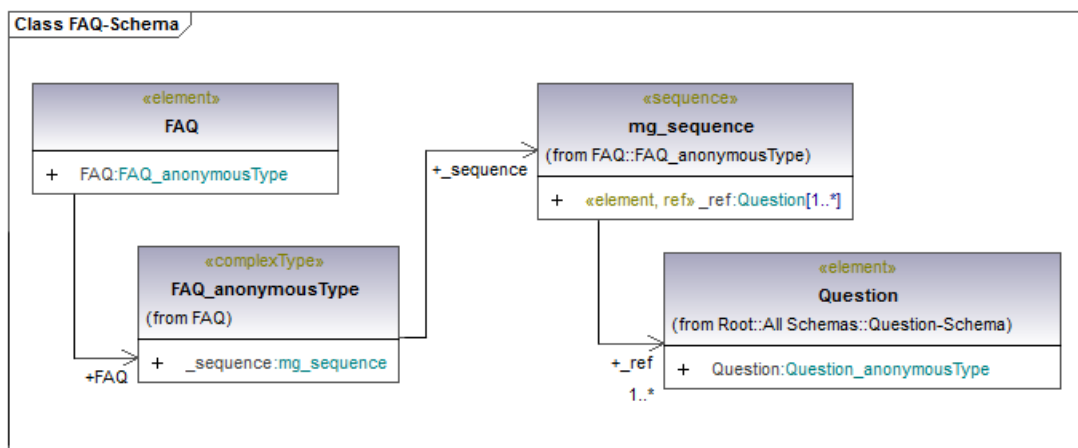


Figura 20 – Modelo de uma FAQ

Uma pergunta, exemplificada na Figura 21, possui vários atributos, um conjunto de conceitos representados por palavras-chave “*Keywords*” e uma interrogativa “*Interrogative*” expressa em várias línguas.

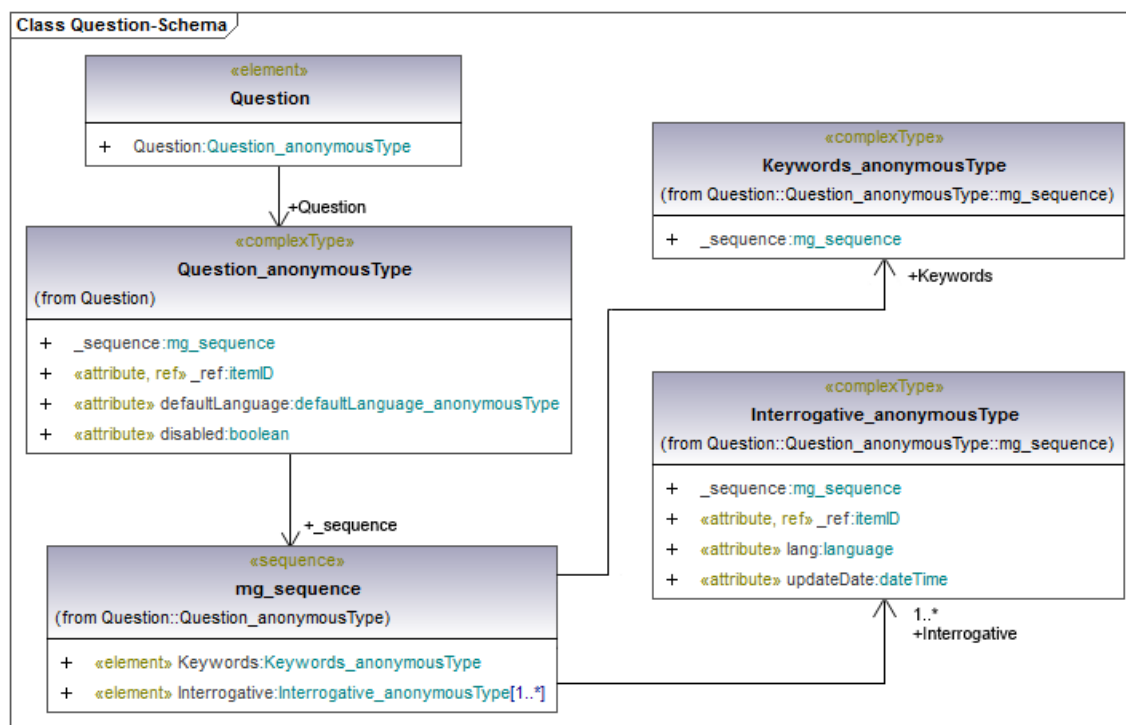


Figura 21 – Modelo de uma Pergunta

Os atributos da pergunta são um identificador único “*ID*”, opcionalmente a língua por omissão “*defaultLanguage*” e finalmente a indicação sobre o estado inativo/ativo da pergunta “*disabled*”.

3.2.2 Palavras-Chave

Uma palavra-chave, indicada na Figura 22, representa um tema ou um conceito do domínio com o qual a pergunta está relacionada.

Além do atributo identificador “*ID*”, um conceito é definido por um descritor linguístico “*LinguisticDescriptor*”, que inclui termos ou expressões associadas, agrupados por língua e um descritor semântico “*SemanticDescriptor*”, constituído por um *Uniform Resource Identifier (URI)* que pode, por exemplo, referir um conceito representado numa ontologia.

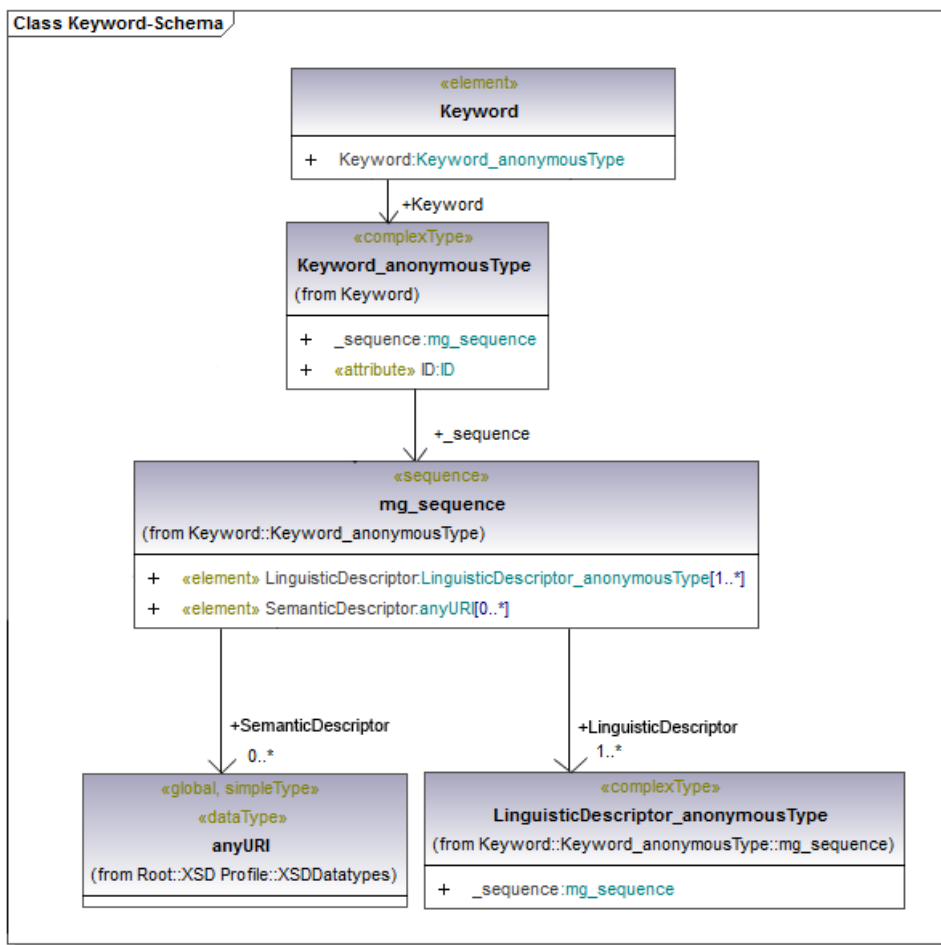


Figura 22 – Modelo de uma *Keyword*

3.2.3 Pergunta

Uma interrogativa, ilustrada na Figura 23, designa o conteúdo de uma pergunta numa determinada língua e inclui o seu texto “*Message*”, uma referência para o utilizador responsável pela criação da pergunta “*Owner*” e um conjunto de pelo menos uma resposta “*Answer*” sendo, por omissão, usada a mais recente.

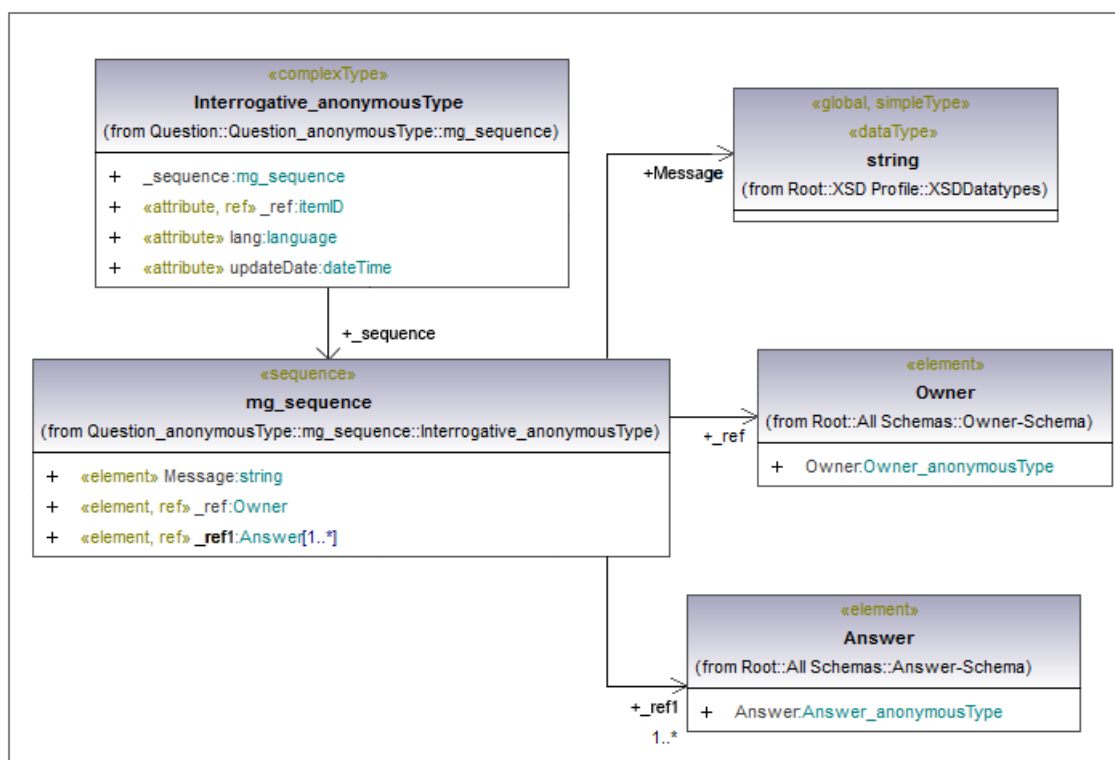


Figura 23 – Modelo da Interrogativa

O facto do corpo da interrogativa suportar várias línguas, permite a definição de perguntas e aquisição de respostas multilíngue sem ter de recorrer a mecanismos externos de tradução, tornando assim possível, por exemplo, quando a fonte é a *Wikipedia* [42] obter a resposta em português no domínio “.pt” ou obter a sua variante inglesa no domínio “.en”.

3.2.4 Resposta

Devido à dimensão do modelo que representa uma resposta optou-se por o dividir em três partes. Na Figura 24 está representada a primeira parte que inclui os atributos da resposta. Além do identificador “*itemID*”, a resposta tem como atributos o seu tipo “*type*”, que indica o formato de apresentação, a data da última atualização “*updateDate*”, o período de revisão “*revisionPeriod*”, a data da última revisão “*revisionDate*”, a data de validade “*expirationDate*” e finalmente o estado inativo/ativo da resposta “*disable*”.

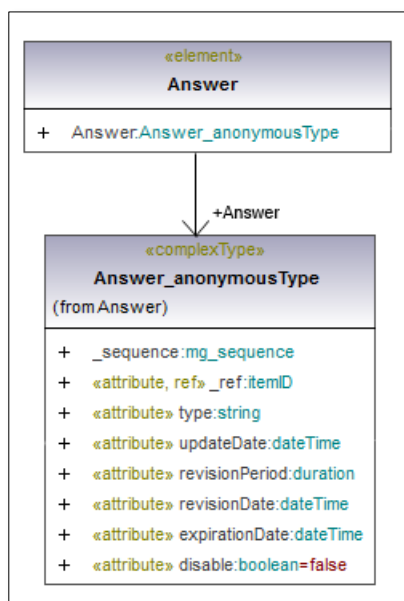


Figura 24 – Atributos de uma Resposta

A segunda parte do modelo da resposta, ilustrada na Figura 25, está ligada à primeira parte através do elemento “*Answer_anonymousType*”, que suporta os classificadores associados à resposta, nomeadamente: as componentes sociais “*Like*” e “*Unlike*”, a origem da resposta “*Source*” e o utilizador que a classificou “*Owner*”.

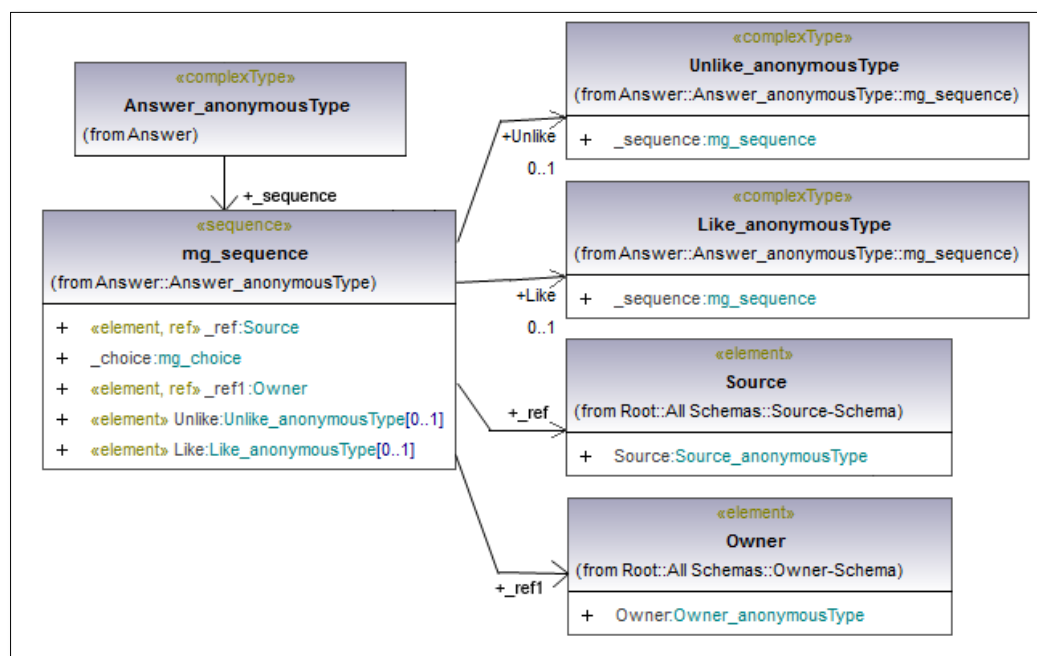


Figura 25 – Classificadores de uma Resposta

A terceira parte do modelo descreve o corpo da resposta que é ilustrado na Figura 26. Esta parte liga-se à segunda parte através da sequência “*mg_sequence*”.

A resposta é obtida de uma fonte de informação, na qual poderá ser necessário a passagem de argumentos. Na eventualidade de não existir nenhuma fonte disponível, a resposta poderá ser armazenada localmente, como uma cache sem prazo de validade.

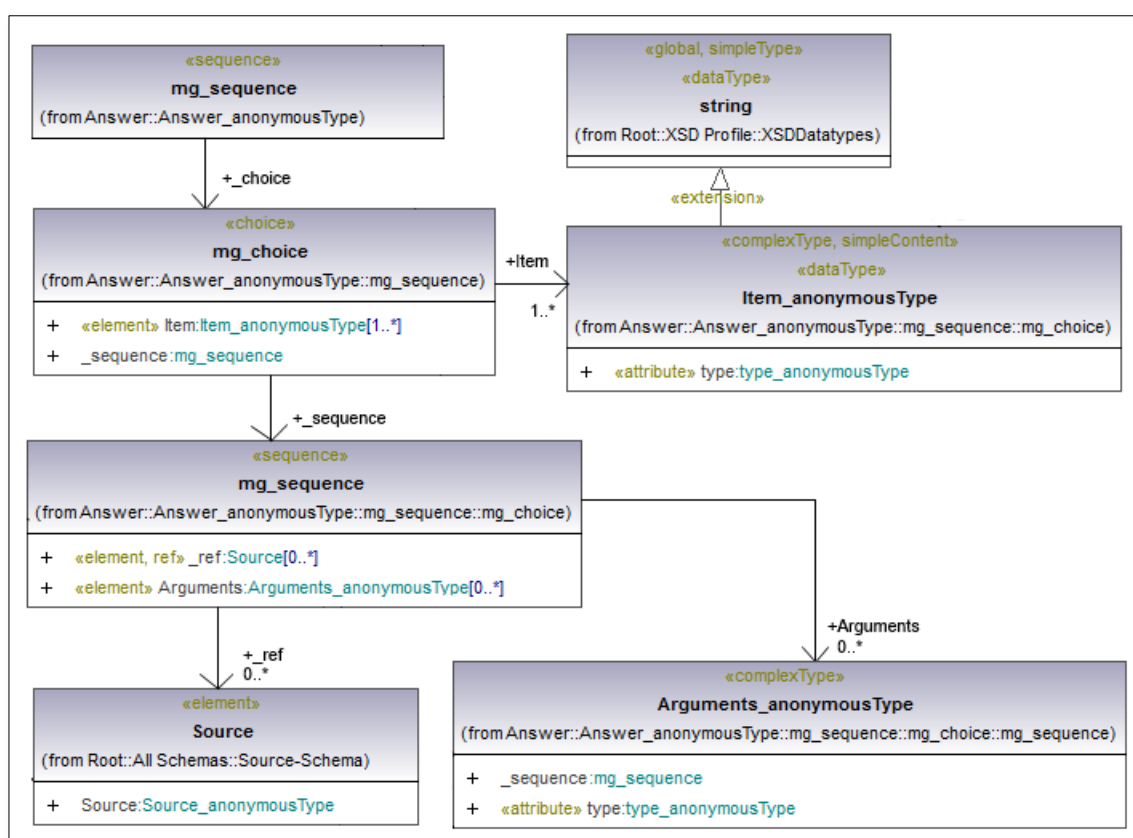


Figura 26 – Elementos de uma Resposta

As fontes de informação são acedidas em tempo de apresentação da resposta e poderão necessitar de argumentos “*Arguments*”. Essas fontes podem ser acedidas via serviços *web* (já existentes) ou por *software* desenvolvido à medida para o efeito. A componente local, designados por “*Item*”, são elementos multimédia definidos por um tipo particular podendo ser, por exemplo, textos, imagens ou vídeos. Dada a sua natureza genérica, é possível fazer combinações destes elementos de forma a modelar uma resposta com um bloco de texto, seguido por uma imagem e novamente por um bloco de texto.

3.2.5 Fonte

A fonte de informação, ilustrada na Figura 27, tem como atributos um identificador “*itemID*”, o tipo de fonte a ser consultada (serviço, RSS, URL) “*type*” e um nome “*name*” assim como um URI para localizar da resposta.

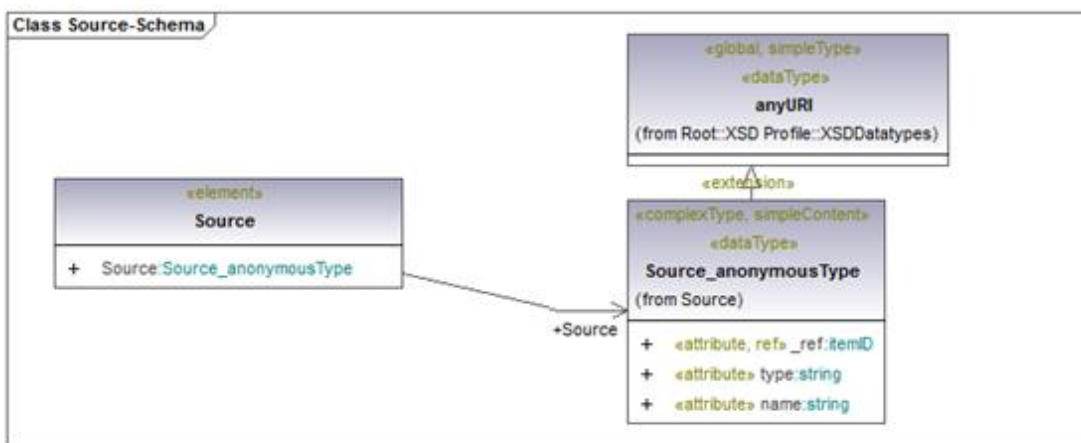


Figura 27 – Modelo da Fonte de Informação

As fontes de informação existem por dois motivos, para corroborar uma resposta fornecida por um utilizador ou para identificar o local de aquisição da resposta.

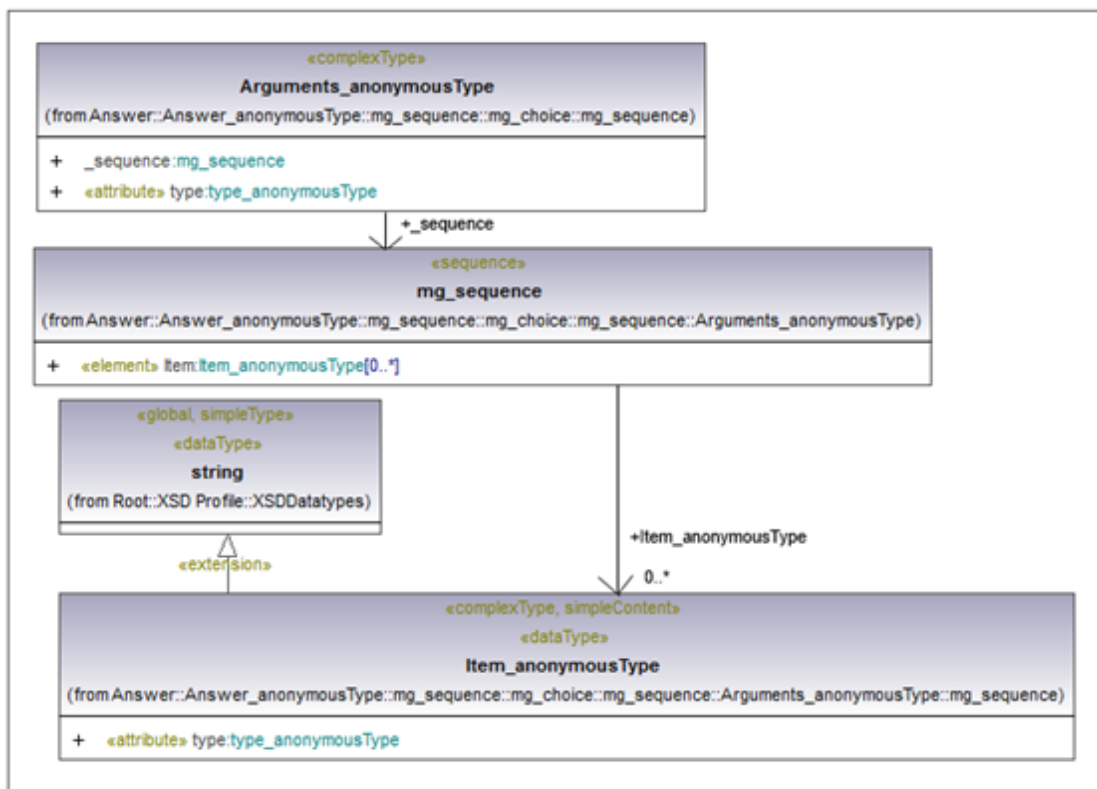


Figura 28 – Modelo dos Argumentos

Para acesso à fonte de informação podem ser indicados argumentos, ilustrados na Figura 28, os quais são de um determinado tipo “*type*”, que pode ser usado para não só especificar o tipo de argumento, como também para agrupar e definir uma ordem dos argumentos passados no pedido efetuado ao serviço. Por exemplo, pode ser necessário passar num serviço valores de duas listas de argumentos diferentes, tais como uma cidade e um meio de transporte, enquanto num outro serviço pode ser passado um argumentos sob forma de texto livre.

3.2.6 Opinião do Utilizador

A opinião do utilizador é modelada recorrendo ao uso dos indicadores “*Like*” e “*Unlike*”, ilustrados respetivamente em cima e em baixo na Figura 29. Estes indicadores são considerados como votos que exprimem o grau de satisfação dos utilizadores.

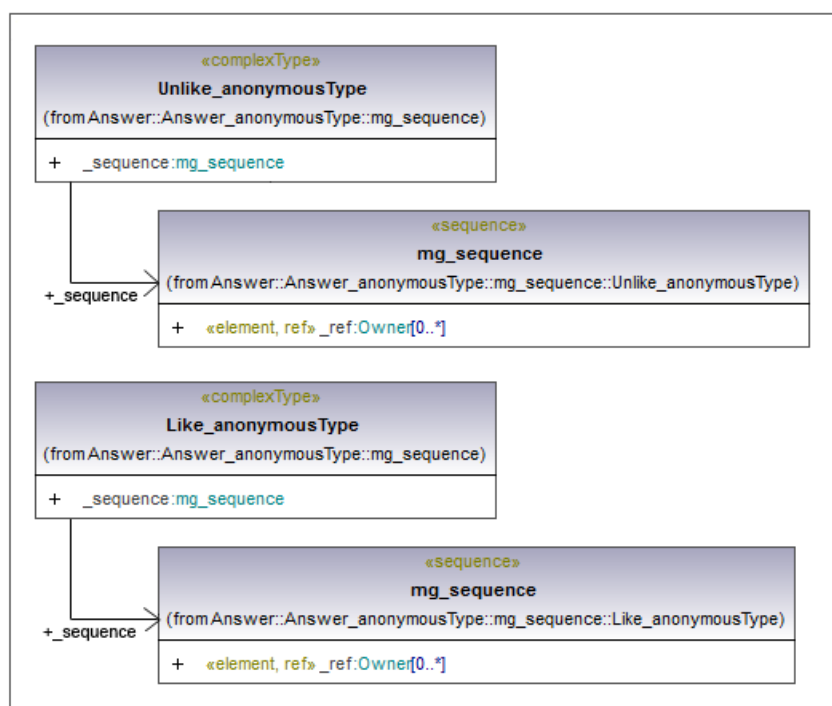


Figura 29 – Modelo do Sistema de Votação

Ambos os indicadores são idênticos porque contêm apenas o utilizador autor do voto. No entanto, para impedir que o mesmo utilizador vote mais do que uma vez é imposta uma restrição.

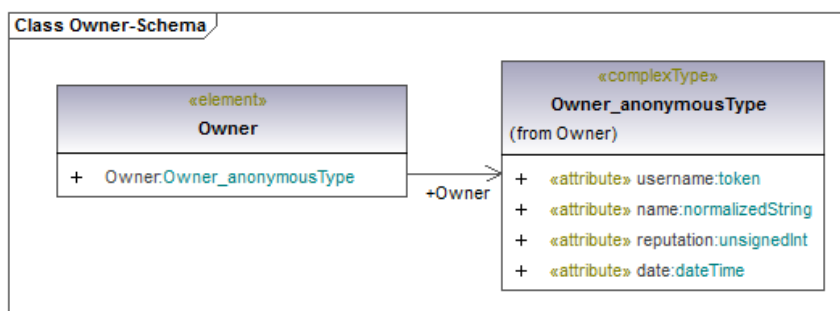


Figura 30 – Modelo do Utilizador

Um utilizador associado a um elemento “*Owner*” é descrito pelo seu nome curto, vulgo nome de utilizador “*username*”, nome “*name*”, reputação “*reputation*” e data da associação à entidade corrente.

3.3 Representação de Informação Prática

A ligação entre uma entidade com a informação prática é feita por um elemento ilustrado no Exemplo 1.

```
<Relation ID="I1" name="CarrisPT" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <Entity type="operator">
    <Operator><Name>Carris</Name> <Country>Portugal</Country></Operator>
  </Entity>
  <PracticalInformation>(…)</PracticalInformation>
</Relation>
```

Exemplo 1 – Relação de uma Entidade com a Informação Prática

Neste elemento, é definida a entidade *Carris* que é ligada à informação prática sendo posteriormente exposta no Exemplo 2.

```
<PracticalInformation xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <Header type="text">Perguntas Frequentes sobre a Carris</Header>
  <Body>(…)</Body>
  <Footer type="image">http://www.carris.pt/temas/default/images/logo.jpg</Footer>
</PracticalInformation>
```

Exemplo 2 – Representação de Informação Prática

A informação prática é constituída pelo cabeçalho, neste caso uma frase introdutória, o corpo e terminado pelo rodapé que é neste caso uma imagem do logótipo do operador.

3.3.1 Representação de Perguntas

No Exemplo 3 está apresentada uma pergunta frequente aplicada a um caso prático, os autocarros da Carris, juntamente com os conceitos associados.

Questão	Palavras-Chave
Como reconhecer um autocarro da Carris?	Transportes; Autocarro

Exemplo 3 – Pergunta e Respetivas Palavras-Chave

A representação desta pergunta é feita no Exemplo 4, através da junção das palavras-chave, a parte interrogativa e a fonte de onde a resposta pode ser recolhida. Também foi associado um utilizador.

```
<FAQ xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <Question itemID="Q6c95fde3" defaultLanguage="pt-PT" disabled="false">
    <Keywords>(...)</Keywords>
    <Interrogative itemID="I75dee158" lang="pt-PT" updateDate="2013-03-
20T23:15:02.1913923+00:00">
      <Message>Como reconhecer um autocarro da Carris?</Message>
      <Owner username="Zé" name="Zé Ninguém" reputation="20" date="2013-03-
20T23:15:02.1913923+00:00"/>
      <Answer>(...)</Answer>
    </Interrogative>
  </Question>
</FAQ>
```

Exemplo 4 – Representação de uma FAQ

As palavras-chave são representadas como um conjunto de termos, como mostra o Exemplo 5.

```
<Keyword ID="ID_1" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <LinguisticDescriptor>
    <CompositeTerm language="pt-PT">
      <Term>Transportes</Term>
    </CompositeTerm>
  </LinguisticDescriptor>
</Keyword>
<Keyword ID="ID_2" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <LinguisticDescriptor>
    <CompositeTerm language="pt-PT">
      <Term>Autocarro</Term>
    </CompositeTerm>
  </LinguisticDescriptor>
</Keyword>
```

Exemplo 5 – Representação de Palavras-Chave

As palavras-chave têm a possibilidade de ter mais sinónimos associados, ou mesmo um descritor semântico.

3.3.2 Representação de Respostas

A resposta é constituída por dois itens de informação, no caso apresentado no Exemplo 6.

```
<Answer itemID="Aaef7a63c" type="Manual" updateDate="2013-03-20T23:15:16.9298001+00:00" revisionDate="2013-03-20T23:15:16.9298001+00:00">
  <Source name="Carris" type="Image" itemID="Sbff7a63c">http://www.carris.pt</Source>
  <Item type="img">http://www.carris.pt/fotos/editor2/autocarro_standard.jpg</Item >
  <Item type="img">http://www.carris.pt/fotos/editor2/autocarro_a_gas_natural.jpg</Item>
  <Item type="text">Autocarro de cor amarela.</Item>
  <Owner username="Zé" name="Zé Ninguém" reputation="20" date="2013-03-20T23:15:02.1913923+00:00"/>
</Answer>
```

Exemplo 6 – Representação de uma Resposta

Neste exemplo, o utilizador forneceu como fonte a página principal do operador, e como resposta colocou duas imagens e um texto descritivo.

3.4 Interface de Serviços

O serviço FAQ ilustrado na Figura 31 descreve as interações disponibilizadas pelo sistema aos utilizadores e administradores.

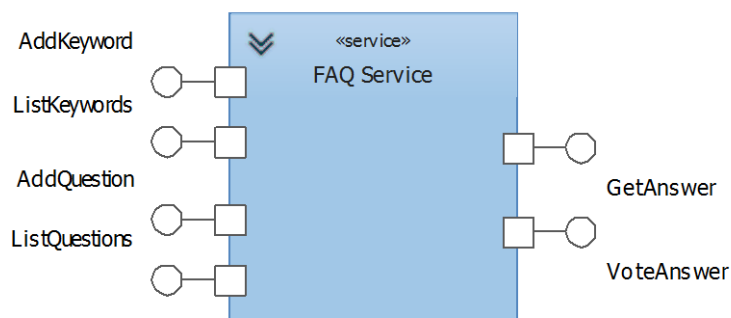


Figura 31 – Interface de Serviços

Estas interações dividem-se em dois grupos, as interações de administração, que contém as operações “AddKeyword”, “ListKeywords”, “AddQuestion” e “ListQuestions”, e as interações de livre utilização, “GetAnswer” e “VoteAnswer”.

As interações de administração são as seguintes:

- a. *AddKeyword* – Adiciona uma palavra-chave ao sistema;
- b. *ListKeywords* – Lista as palavras-chave existentes;
- c. *AddQuestion* – Adiciona uma questão ao sistema;
- d. *ListQuestions* – Lista as questões presentes;

De notar que os retornos das operações de listagem, “*ListKeywords*” e “*ListQuestions*” podem ser filtradas à partida através de Palavras-Chave e/ou línguas.

Já as interações que são usadas por todos os utilizadores são:

- a. *GetAnswer* – Obtém a resposta a uma pergunta específica;
- b. *VoteAnswer* – Atribui o voto do utilizador registado à resposta.

A operação “*VoteAnswer*” apenas pode ser executada perante um utilizador válido, por restrições do modelo.

4 Demonstrador

Neste capítulo é apresentado o demonstrador utilizado para validar a arquitetura proposta. São apresentados exemplos de informação prática sobre transportes modelada sob a forma de perguntas frequentes sobre autocarros.

4.1 Análise Funcional

Foi efetuado um estudo focado na ótica de utilização do sistema sobre as funcionalidades que devem estar presentes no desenvolvimento do demonstrador.

4.1.1 Casos de Utilização

Foram identificadas três interações com o utilizador ilustradas na Figura 32.

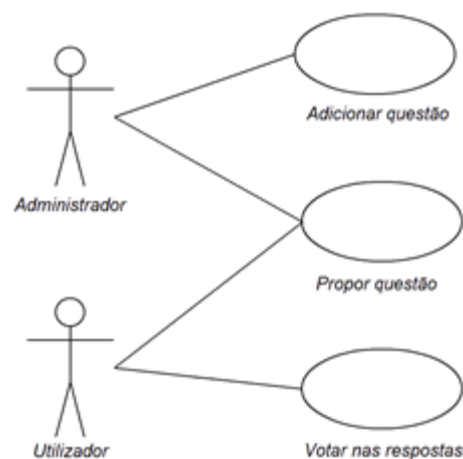


Figura 32 – Casos de Utilização

A adição de perguntas na lista de FAQ é uma ação passível de ser feita apenas por um administrador do portal. As outras duas ações podem ser feitas por qualquer utilizador registado e são respetivamente: propor uma pergunta/resposta e votar nas respostas. O administrador pode eventualmente optar por adicionar a pergunta proposta à lista existente de perguntas.

4.1.2 Diagramas de Sequência

Na Figura 33 e na Figura 34 estão descritas as trocas de mensagens que ocorrem quando a interação de um administrador ou um cliente com o portal.

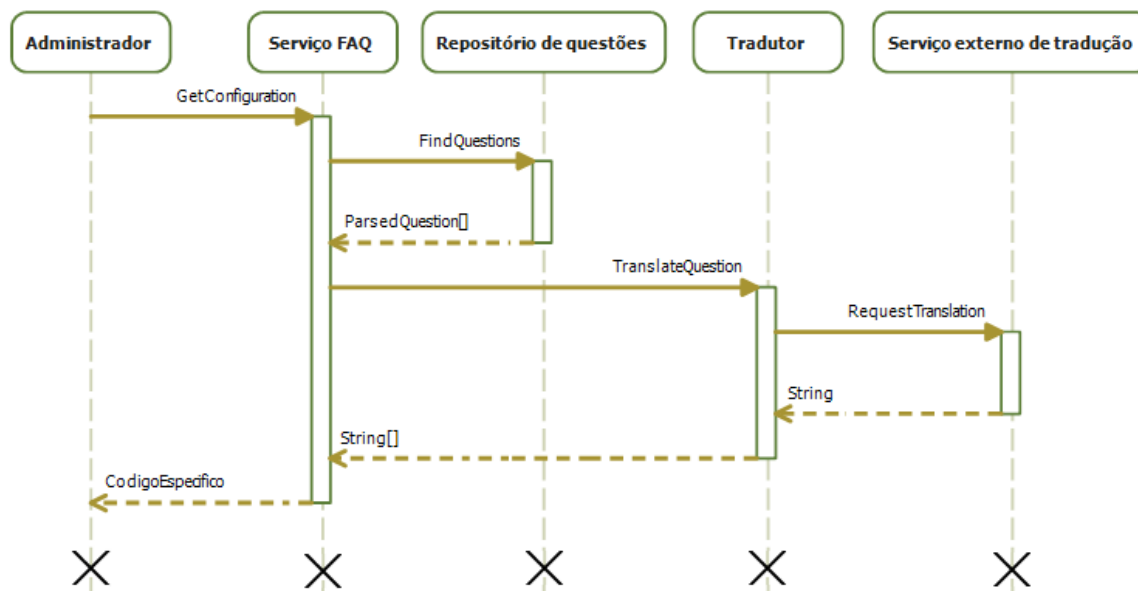


Figura 33 – Diagrama de Sequência para a Configuração de Perguntas

Através do Administrador, o portal pede a lista de perguntas, filtrada através de um conjunto de termos a enviar, uma língua específica e a tecnologia em que as perguntas irão ser representadas. O Serviço FAQ pede então ao repositório de perguntas a lista filtrada por termos e língua, o qual retorna um conjunto de perguntas que obedecem obrigatoriamente ao conjunto de termos, e opcionalmente à língua. As perguntas filtradas são constituídas por um identificador, a língua em que se encontra o conteúdo e o texto da pergunta. Após a receção do conjunto, as perguntas são analisadas para confirmar a língua, e as que não se encontram na língua destino são enviadas para o serviço de tradução. O tradutor, através de mecanismos tais como bases de dados internas, ou recorrendo a outros serviços de tradução, retorna o texto e a língua em que se encontra, que poderá ser o mesmo do envio, caso não tenha tido sucesso. Por fim, o Serviço FAQ constrói sob forma de texto as perguntas na tecnologia pretendida, para poderem ser usadas diretamente no portal. Isto permite que o Administrador não se tenha de preocupar em adaptar a resposta do serviço para mostrar as perguntas, uma vez sabendo que se tiver o portal em HTML, as perguntas virão nesta linguagem.

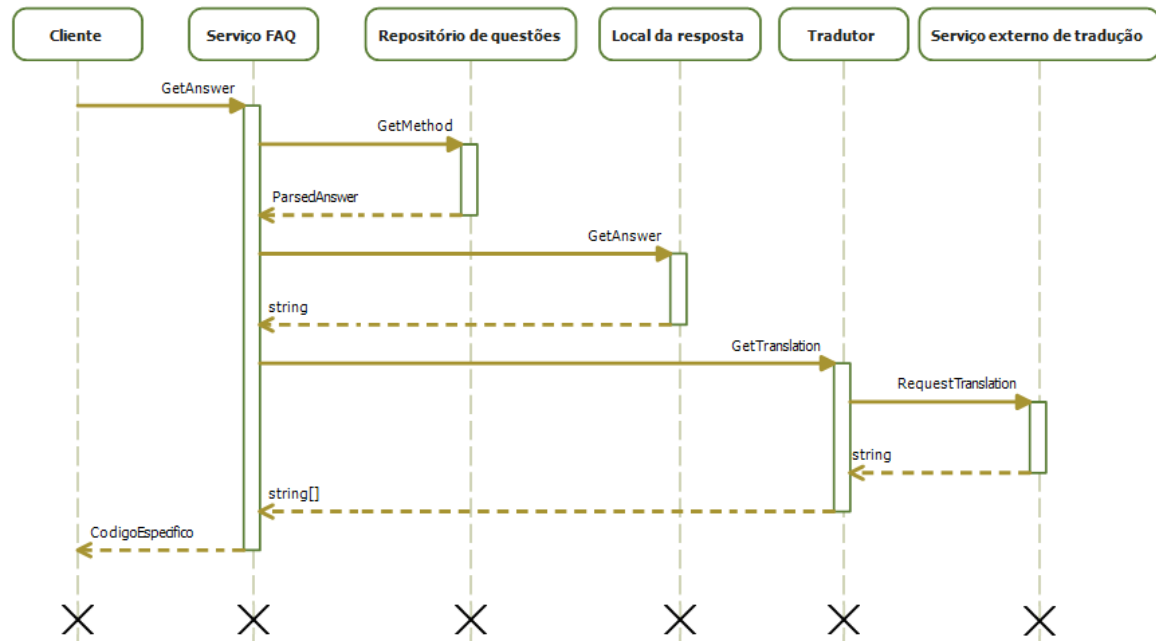


Figura 34 – Diagrama de Sequência para a Obtenção de Respostas

Quando o cliente, normalmente o portal, pede a resposta a uma pergunta, através da passagem do identificador desta, a língua e a tecnologia em que pretende a resposta, o serviço contacta o repositório para obter o método de resposta. Este método pode ser, por exemplo, texto armazenado localmente, uma imagem, o contacto com outro serviço idêntico ou uma hiperligação para uma página ou um planeador. Uma vez obtido o método, conforme o tipo deste é efetuada a ação apropriada, e caso seja obtida uma resposta textual, é feita a validação da língua, e por fim a transformação do texto na tecnologia apropriada, tal como descrito no caso anterior.

4.2 Implementação do Módulo de FAQ

A aplicação do módulo de FAQ é uma ferramenta de administração através da qual são adicionadas perguntas ao sistema cliente, utilizando o serviço que implementa a arquitetura até agora descrita. Na Figura 35 ilustra-se a página visível quando o sistema é posto em execução a primeira vez.

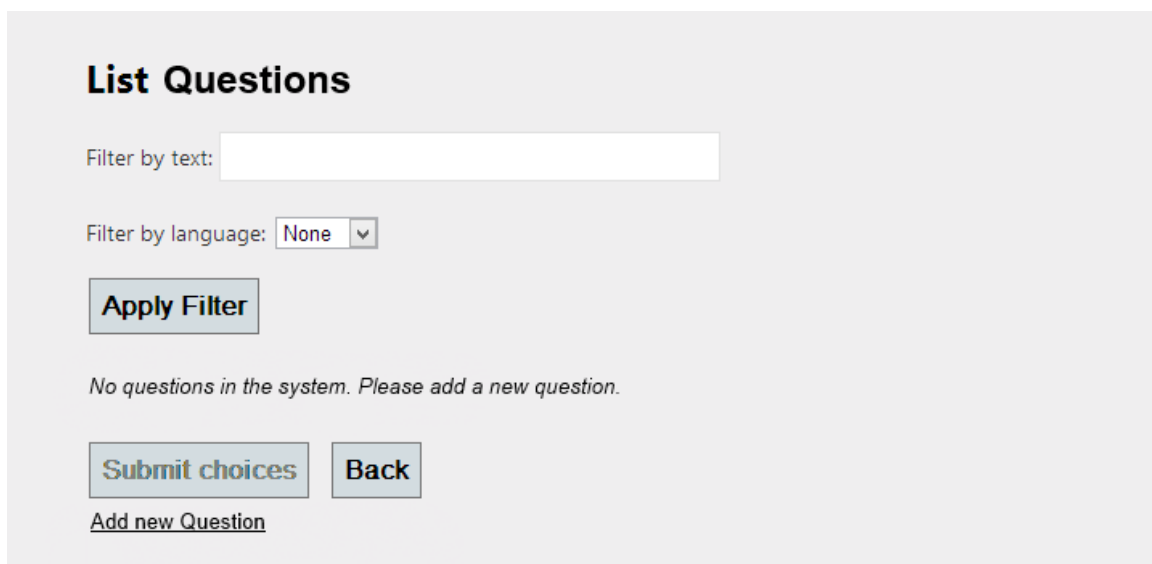


Figura 35 - Lista de Perguntas Vazia

Nesta situação, como não existe ainda qualquer pergunta, inicia-se o processo através da ligação “Add new Question”. A definição de uma pergunta passa por várias etapas, desde a escrita do texto da pergunta, apresentada na Figura 36, à atribuição de uma resposta e associação a palavras-chave.

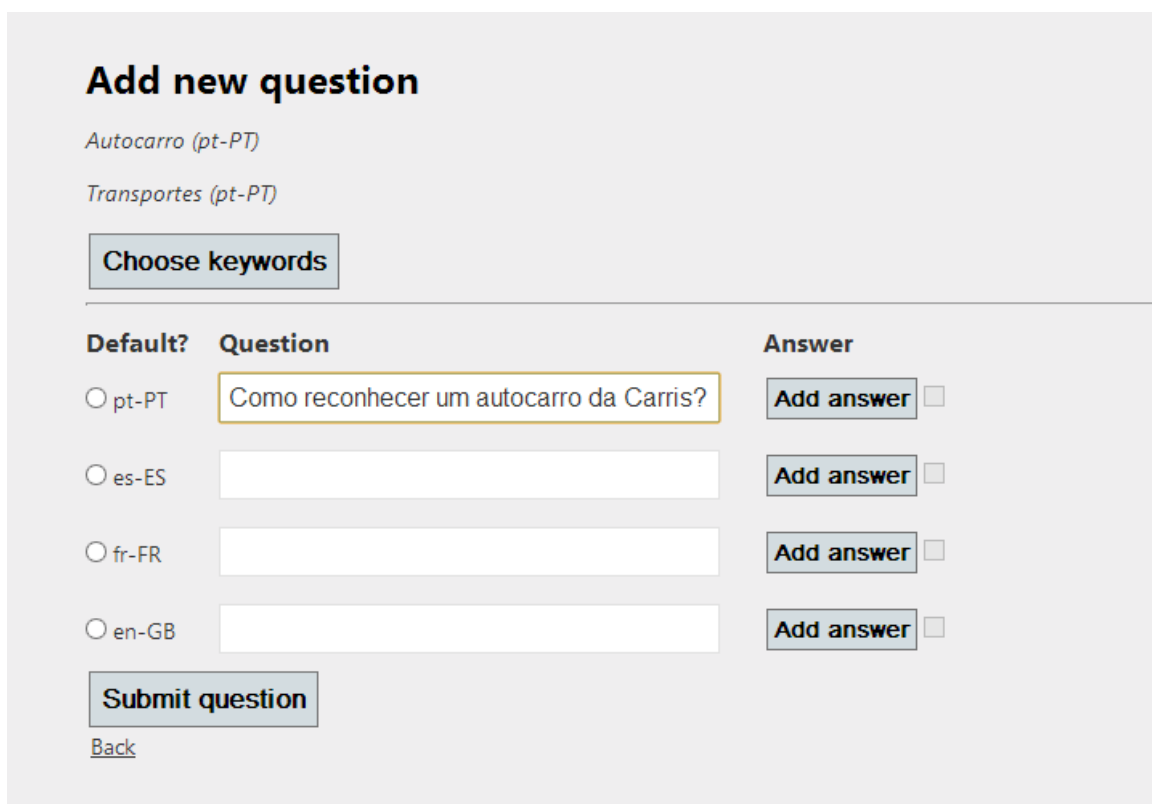


Figura 36 - Adição de uma Pergunta

A adição de uma pergunta apresenta a possibilidade de escrita do texto em todas as línguas suportadas pelo sistema, bem como a atribuição de respostas diferentes para cada uma delas. Na figura apenas está a ser inserida a pergunta em português, “Como reconhecer um autocarro da Carris?”.

Figura 37 - Adição de uma Resposta

Na Figura 37 representa-se a adição de uma resposta. Neste caso, a resposta é local, e é composta por três itens, duas imagens e um bloco de texto. Após a inserção dos vários itens, a resposta é submetida e associada à pergunta.

Figura 38 - Escolha de Palavras-Chave

Ao associar Palavras-Chave, é apresentada uma lista de palavras-chave “*Keyword*” nas várias línguas, presente na Figura 38.

Caso se pretenda adicionar uma Palavra-Chave inexistente na lista, é possível a adição de novas Palavras-Chave, através do formulário visível na Figura 39.



Add New Keyword

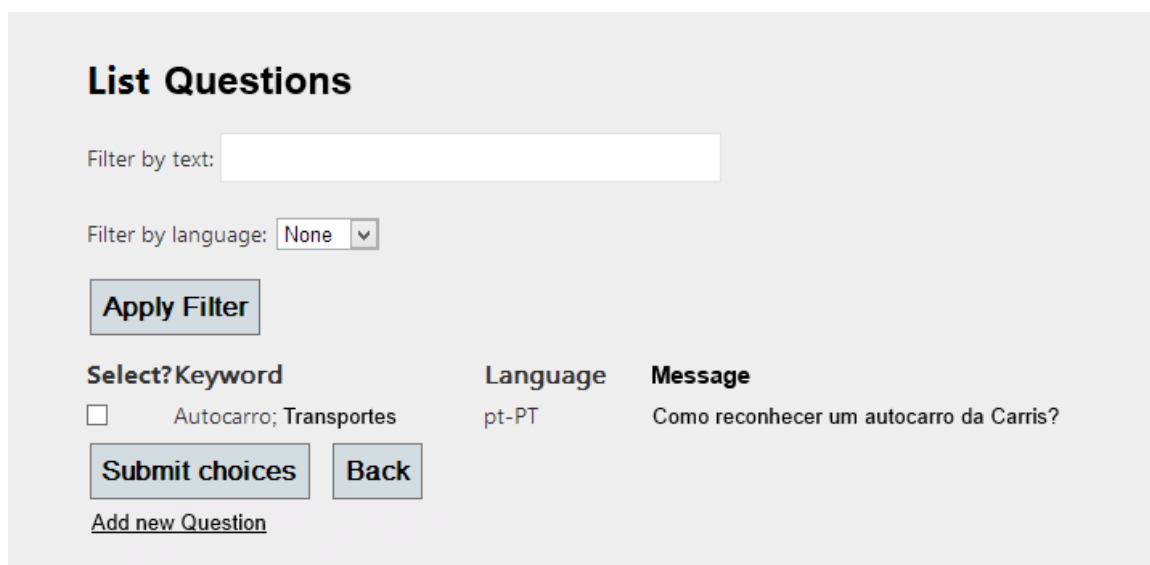
Term Language

[Back](#)

Figura 39 - Adição de uma Palavra-Chave

Este formulário permite a associação de várias palavras “*Term*” em várias línguas, e respetivas variantes. Uma palavra com variação é, por exemplo, a cidade alemã de Aachen, que possui duas variantes na língua portuguesa, nomeadamente Aquisgrão e Aquisgrano.

Após a submissão das Palavras-Chave, é-se retornado para a página de adição da pergunta, permitindo assim a continuação do processo.



List Questions

Filter by text:

Filter by language:

Select? Keyword	Language	Message
<input type="checkbox"/> Autocarro; Transportes	pt-PT	Como reconhecer um autocarro da Carris?

[Add new Question](#)

Figura 40 - Escolha de Perguntas

Aquando da submissão da pergunta, a lista de perguntas é reconstruída e agora a pergunta recém-adicionada aparecerá na lista ilustrada na Figura 40, e permite a sua escolha pelo administrador para incorporar a lista de perguntas do portal.

4.3 Cenários de Teste

Foi modelado um conjunto de perguntas, além das já indicadas na secção 2 deste documento, para validar o demonstrador, descritas na Tabela 5.

Entidade	Palavra-chave	Pergunta	Resposta	
1	Taxi	Tarifas	Qual a tarifa mínima aplicada?	2€ durante a semana e 2,5€ no restante período
2	Taxi	Tarifas	Qual a tarifa média aplicada?	0,45€/km ou 0,90€/km se ida e volta
3	Taxi	Tarifas	Qual o método de pagamento?	Dinheiro ou MB
4	Taxi	Horário	Qual o horário de funcionamento?	24/7
5	Taxi	Acessibilidade	É possível o transporte de cadeiras de rodas?	Sim
6	Taxi	Acessibilidade	É possível o transporte de bagagens?	Sim, com custo adicional
7	Taxi	Acessibilidade	É possível o transporte de animais?	Sim, com custo adicional
8	Taxi	Contactos	Qual o contacto da operadora [XYZ]?	[XYZ] : [xxxxxxxxx]
9	Metro	Tarifas	Qual é o preço de uma viagem?	1,25€
10	Metro	Tarifas	Qual é o preço de um bilhete diário?	5€
11	Metro	Tarifas	Há desconto para crianças?	Sim, menores de 4 anos não pagam bilhete
12	Metro	Acessibilidade	A estação [XYZ] é acessível?	Sim/Não
13	Metro	Acessibilidade	É possível o transporte de bagagens?	Sim
14	Metro	Acessibilidade	É possível o transporte de animais?	Sim, com os devidos cuidados
15	Metro	Acessibilidade	É possível o transporte de bicicletas?	Sim, se houver pouca gente
16	Metro	Horário	Qual o horário de funcionamento?	06:30 – 01:00
17	Metro	Diagrama	Onde posso encontrar um diagrama da rede?	<i>Imagem</i>
18	Bicicleta	Aluguer	Onde posso alugar uma bicicleta?	<i>URL</i>
19	Bicicleta	Transporte	Posso transportar uma bicicleta no transporte [XYZ]?	Sim/Não
20	Bicicleta	Mapa	Onde posso encontrar um mapa das ciclovias?	<i>URL</i>

Tabela 5 – Perguntas Modeladas

As perguntas modeladas têm, na sua maioria, respostas simples e diretas e são reutilizáveis. No entanto, as perguntas configuráveis, nomeadamente as

indicadas pelos números 8, 12 e 19 necessitam de ter argumentos configurados, para permitir respostas dinâmicas.

4.4 Adaptação à *Integra Social Network*

A ISN é um sistema computacional que implementa uma rede social temática relacionada com a mobilidade sustentável. No contexto da ISN, são partilhadas experiências de mobilidade (viagens) que são reutilizadas para fazer recomendações.

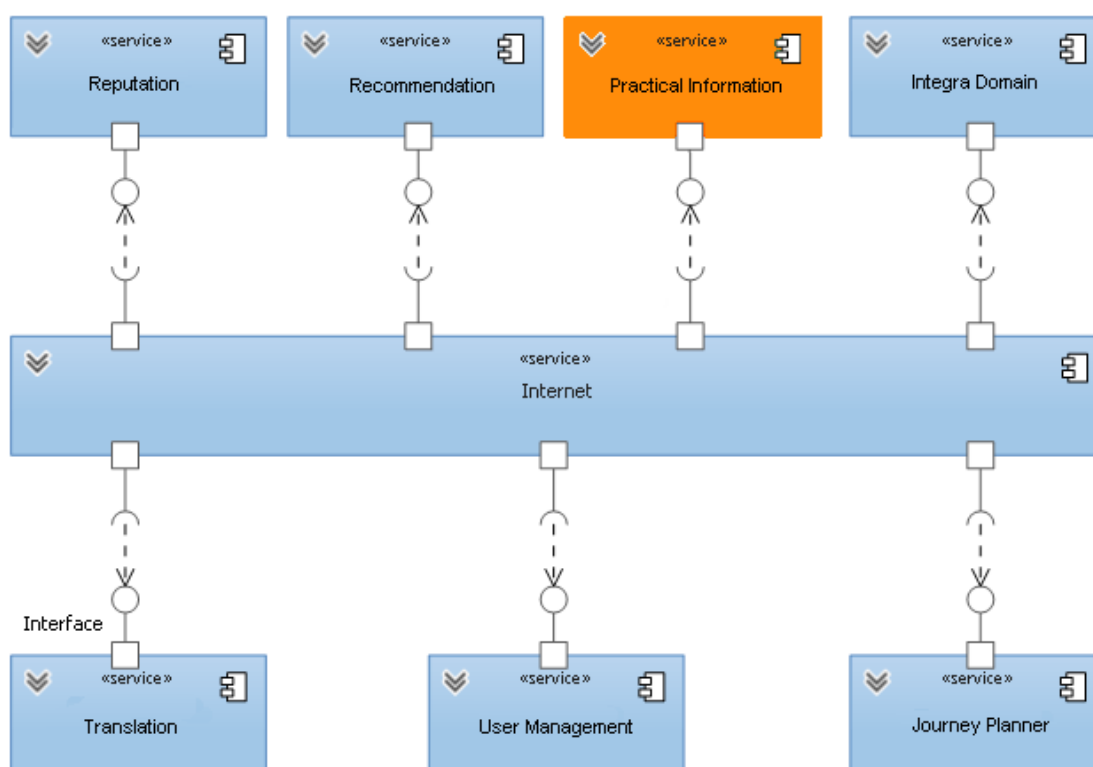


Figura 41 – Arquitetura SOA da ISN

A Figura 41 descreve a arquitetura de serviços da ISN onde o serviço de Informação Prática proposto neste trabalho aparece representado. Os serviços existentes descrevem-se genericamente, como:

- *Reputation*: utilizado para calcular a reputação dos utilizadores;
- *Recommendation*: faz recomendações com base nos gostos dos utilizadores;
- *User Management*: gestão dos perfis dos utilizadores;
- *Translation*: usado para obter as várias traduções;
- *Journey Planner*: propõe viagens multimodais entre dois pontos;

- *Integra Domain*: orquestra todos os serviços acima descritos.

O serviço de informação prática proposto interage com os serviços existentes, principalmente com a representação do domínio, adicionalmente também interage com serviços externos necessários para a aquisição das respostas.

4.5 Simulação do Módulo na *Integra Social Network*

A título de exemplo apresentam-se questões na Língua Portuguesa, presentes na Tabela 6, com diversas palavras-chave e sem associação inicial ao módulo.

ID	Perguntas do sistema	Palavras-chave
1	Onde comprar e validar o bilhete?	Portugal; Transportes; Comboio; CP
2	Como reconhecer um autocarro da Carris?	Lisboa; Transportes; Autocarro; Carris
3	Como se entra e saí do metro?	Lisboa; Transportes; Metro
4	Qual língua da informação a bordo?	Portugal; Transportes; Comboio; CP
5	Como assinalar ao condutor de um autocarro que deve parar?	Porto; Transportes; Autocarro

Tabela 6 – Perguntas Exemplo

Cabe primeiramente a um administrador ou a um utilizador com maiores privilégios adicionar primeiramente as perguntas. Para esse efeito, acede ao menu FAQ, neste exemplo da página da operadora Comboios de Portugal (**CP**), como apresentado na Figura 42.

Nesta situação, o Administrador acedeu ao motor de pesquisa com o valor CP, e já assinalou ambas as perguntas que apareceram disponíveis, em inglês.

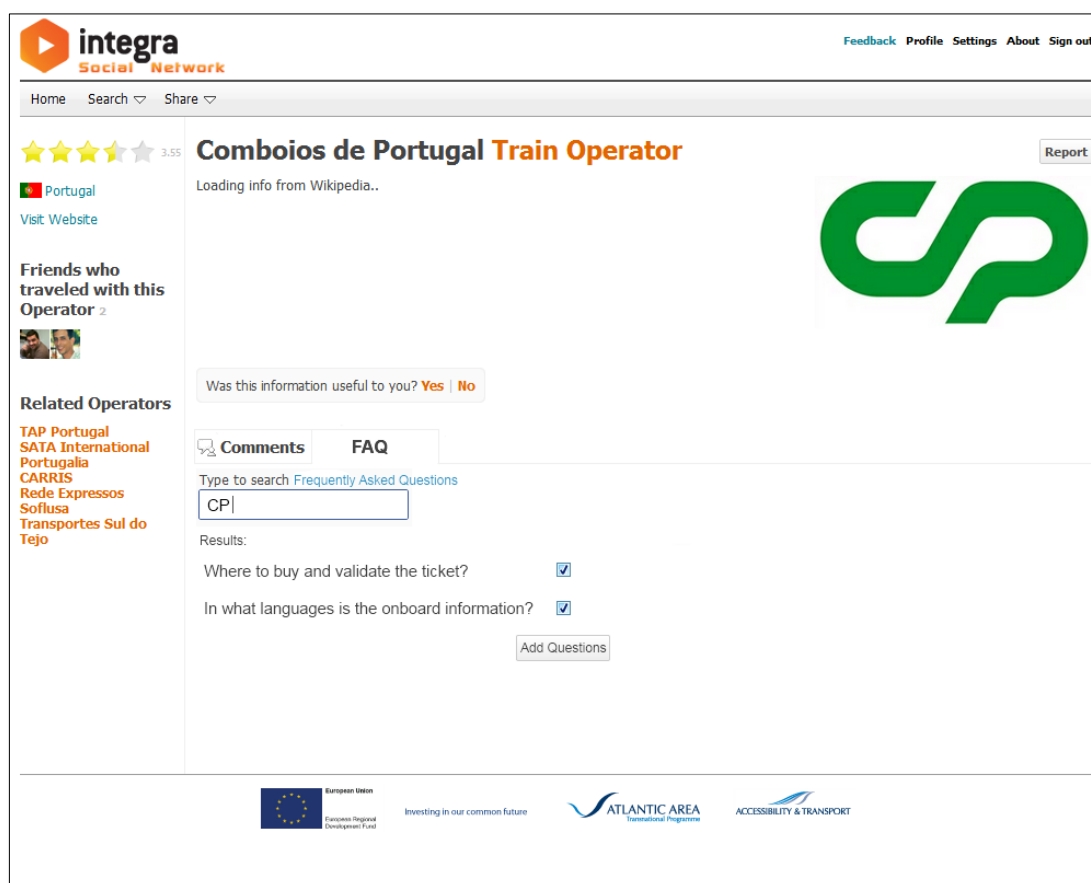


Figura 42 – Configuração de Perguntas Frequentes

Nos bastidores, foi feito um pedido ao Serviço FAQ para listar as perguntas com os seguintes dados: “CP”, “en” e “html”. Ao receber o pedido, o serviço reencaminha o conjunto de termos para um dicionário de sinónimos, o qual retorna o mesmo valor. Após este passo, os dois primeiros parâmetros são enviados para o serviço de aquisição de informação. Neste caso, como não existem perguntas em inglês, o serviço de dados retorna o texto em português de ambas as perguntas e respetivos identificadores. Ao obter a lista, é feita a validação para a língua, e neste caso ambas são reencaminhadas para o serviço de tradução. Este serviço percorre uma base de dados, obtém a tradução, atualiza as perguntas e reenvia-as para o Serviço FAQ. Por fim, o Serviço FAQ transforma as perguntas já tratadas numa *string* em HTML e envia para o cliente.

The screenshot shows the 'integra Social Network' interface. At the top, there are navigation links: Home, Search, and Share. The main content area displays a search result for 'Comboios de Portugal Train Operator' with a 3.55 star rating and a 'Report' button. Below the title, it says 'Loading info from Wikipedia..'. To the right is a large green logo. A feedback prompt asks 'Was this information useful to you? Yes | No'. Below that are tabs for 'Comments', 'FAQ', and 'Add Questions'. The FAQ section contains two questions: 'Where to buy and validate the ticket?' and 'In what languages is the onboard information?'. The sidebar on the left includes 'Friends who traveled with this Operator' and 'Related Operators' such as TAP Portugal, SATA International, CARRIS, Rede Expressos, Soflusa, and Transportes Sul do Tejo. The footer features logos for the European Union, Atlantic Area, and Accessibility & Transport.

Figura 43 – Visualização de Perguntas Frequentes

Ao aceder à mesma página, como pode ser visto na Figura 43, um utilizador sem privilégios de administração visualiza as perguntas adicionadas anteriormente, mas já respondidas.

Nesta situação, à semelhança do primeiro pedido para adicionar as perguntas, foram feitos vários pedidos para responder às várias questões, tendo sido passados os identificadores das perguntas 1 e 4, bem como em ambos os casos os parâmetros “en” e “html”. Mais uma vez, um pedido é feito para o Serviço FAQ, e este por sua vez pede ao serviço de dados para obter a resposta às perguntas em inglês. O serviço de dados verifica que as perguntas têm as respostas localmente, mas em português, e tal como na obtenção das perguntas, retorna os dados existentes. O resto do processamento é igual, com a diferença do tipo de objeto entre os vários serviços, pois contém mais dados, nomeadamente a fonte da informação e as datas de validade. No fim, é gerado o código para o cliente e retornado.

5 Epílogo

Neste capítulo é feito um balanço do trabalho realizado sendo posteriormente apresentada a conclusão e perspectivas de trabalho futuro.

5.1 Balanço

No âmbito deste trabalho, foi proposta uma arquitetura para adquirir informação prática sobre entidades, pertencentes a um determinado domínio, acedida por um conjunto de perguntas frequentes usado para adquirir as respetivas respostas. Para efeitos de avaliação experimental, foi implementado um demonstrador no domínio da mobilidade sustentável que adaptado à *Integra Social Network*.

No âmbito do domínio Integra considerou-se a existência de um conjunto de perguntas frequentes sobre, por exemplo, operadores, meios de transporte sustentáveis, locais e eventos. As perguntas são baseadas no conceito 5W com o objetivo de transformar a informação, sobre uma determinada temática, num conjunto de perguntas/respostas claras e concisas.

A aquisição de informação é concretizada com recurso à técnica *Mashup*, que consiste em recolher informação de fontes heterogéneas, normalizá-la num formato comum, para ser utilizada de forma transparente e independente da sua fonte. Esta técnica assenta numa arquitetura orientada a serviços, com orquestração que conduz à flexibilidade e extensibilidade.

5.2 Conclusão

As atividades desenvolvidas no âmbito deste trabalho consistiram em descrever e operacionalizar o processo de integração dinâmica de informação, oriunda de diversas fontes heterogéneas e respetiva contextualização numa rede social. Neste âmbito, foram concretizados os objetivos, nomeadamente:

- a. Descrever a arquitetura para aquisição de respostas;
- b. Descrever o modelo para representar a informação prática;
- c. Implementar um serviço pronto para produção;
- d. Adaptar a arquitetura proposta à *Integra Social Network*.

A validação experimental da arquitetura proposta foi realizada com sucesso através da implementação de um demonstrador adaptado à *Integra Social Network*. A conceção da arquitetura proposta é orientada a serviços com o objetivo de facilitar a interoperabilidade entre os vários protagonistas intervenientes na definição do domínio tal como, por exemplo, a mobilidade sustentável.

Por último, salienta-se que foi apresentado numa conferência um artigo que divulga as ideias base do trabalho realizado [40].

5.3 Perspetivas de Trabalho Futuro

No decorrer deste trabalho foram identificados alguns aspetos que podem ser melhorados, designadamente:

- a. A melhor integração entre o demonstrador e a *Integra Social Network* no que respeita às múltiplas línguas;
- b. Composição tendencialmente automática de serviços para a aquisição de respostas.

Referências

- [1] *Official Journal of the European Union*, 2009 - *DIRECTIVE 2009/28/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL* de 23 de abril de 2009
- [2] *Merseytravel* através do Consórcio START. 2009 - 2011. *START Project - Objectives*. [Online] Disponível em: <http://www.start-project.eu/en/Objectives.aspx> [Acedido em 22 de agosto de 2012]
- [3] Carris – Transportes Públicos Lisboa, 2011. *Sustainability*. [Online] Disponível em: <http://www.carris.pt/pt/sustentabilidade/> [Acedido em 20 de agosto de 2012]
- [4] Wang, R., Strong, D. (1996) *Beyond Accuracy: What Data Quality Means to Data Consumers*. *Journal of Management Information Systems*, 12 (4), p. 5-34.
- [5] Carris – Transportes Públicos Lisboa, 2011. *Carris – Transportes Públicos Lisboa*. [Online] Disponível em: <http://www.carris.pt/> [Acedido em 20 de agosto de 2012]
- [6] Metropolitano de Lisboa, E.P.E, 2012. [Online] Disponível em: <http://www.metrolisboa.pt/> [Acedido em 20 de agosto de 2012]
- [7] CP - Comboios de Portugal. *CP Passageiros :: CP :: (Versão Portuguesa)*. [Online] Disponível em: <http://www.cp.pt/> [Acedido em 21 de agosto de 2012]
- [8] Domingues, D., Félix, A., Aires, R., 2010. *SITP – Sistema de Informação sobre Transportes Públicos* [Online] Disponível em: <http://start.isel.pt/sitp/> [Acedido em 3 de julho de 2012]
- [9] *Commentjyvais.fr* [Online] Disponível em: <http://commentjyvais.fr> [Acedido em 4 de julho de 2012]
- [10] *BreizhGo* [Online] Disponível em: <http://www.breizhgo.com> [Acedido em 4 de julho de 2012]

- [11] GuIAA – Grupo de Investigação em Ambientes Autónomos. 2009. *Atlantic Area START related projects* [Online] Disponível em: <http://start.isel.pt/> [Acedido em 2 de julho de 2012]
- [12] TRANSPORLIS, 2009. *Transporlis > Home* [Online] Disponível em: <http://www.transporlis.sapo.pt/> [Acedido em 2 de setembro de 2012]
- [13] Costa, A., Antunes, D., 2012. *Integra Social Network*. [Online] Disponível em: <http://integra.isel.pt> [Acedido em 20 de abril de 2012]
- [14] ISEL – Instituto Superior de Engenharia de Lisboa, 2012. *ISEL – Instituto Superior de Engenharia de Lisboa* [Online] Disponível em: <http://www.isel.pt/> [Acedido em 25 de setembro de 2012]
- [15] *Merseytravel* através do Consórcio START. 2009 - 2011. *START Project - Home*. [Online] Disponível em: <http://www.start-project.eu/> [Acedido em 22 de agosto de 2012]
- [16] Google, 2012. *Google* [Online] Disponível em: <http://www.google.com/> [Acedido em 25 de setembro de 2012]
- [17] Microsoft, 2012. *Bing* [Online] Disponível em: <http://www.bing.com/> [Acedido em 25 de setembro de 2012]
- [18] Kuchmann-Beauger, N., Aaufaure, M. *Structured data-based Q&A System using surfasse patterns*. Em atas da 9^a *International Conference of Flexible Query Answering Systems (FQAS)*. 2011. Ghent, Bélgica.
- [19] Rosso, P., Benajiba, Y. *Towards na Arabic Question Answering System*. Em atas da 4^a *Scientific Research Outlook & Technology Development in the Arab World (SROIV)*. Dezembro de 2006. Damasco, Síria.
- [20] Katz, B. *The START Natural Language Question Answering System* [Online] Disponível em: <http://start.csail.mit.edu/> [Acedido em 17 de maio de 2012]
- [21] NLX-Grupo de Fala e Linguagem Natural da Universidade de Lisboa, Departamento de Informática. *XisQuê – Resposta a perguntas na Web portuguesa* [Online] Disponível em: <http://xisque.di.fc.ul.pt/> [Acedido em 17 de maio de 2012]
- [22] *Department for Transport*, Governo Britânico, *Identification of Fixed Objects in Public Transport (IFOPT)* [Online] Disponível em: <http://www.dft.gov.uk/naptan/ifo/ifoV1.0->

- [36/CENTC278WG3SG6_IFOPT_20081110_36.pdf](#) [Acedido em 4 de agosto de 2012]
- [23] Wasserman, S., Faust, K., 1994. *Social Network Analysis in the Social and Behavioral Sciences. Social Network Analysis: Methods and Applications*. Cambridge University Press. pp. 1–27. ISBN 978521382694.
- [24] Boyd, D. M., Ellison, N. B. (2007). *Social Network Sites: Definition, History, and Scholarship*. *Journal of Computer-Mediated Communication*, 13(1), artigo 11.
- [25] hi5 Networks, 2012. *hi5* [Online] Disponível em: <http://www.hi5.com> [Acedido em 1 de maio de 2012]
- [26] LinkedIn Corporation, 2013. *LinkedIn* [Online] Disponível em: <http://www.linkedin.com> [Acedido em 1 de maio de 2012]
- [27] Google, 2012. *Orkut* [Online] Disponível em: <http://www.orkut.com> [Acedido em 1 de maio de 2012]
- [28] Twitter, 2012. *Twitter*. [Online] Disponível em: <https://www.twitter.com> [Acedido em 1 de maio de 2012]
- [29] MySpace LLC, 2003-2012. *MySpace* [Online] Disponível em: <http://www.myspace.com> [Acedido em 1 de maio de 2012]
- [30] Facebook, 2012. *Facebook*. [Online] Disponível em: <http://www.facebook.com> [Acedido em 1 de maio de 2012]
- [31] Google, 2012. *Google+*. [Online] Disponível em: <https://plus.google.com> [Acedido em 1 de maio de 2012]
- [32] ComScore – *Google+ Grows Worldwide Users From 65 Million In October To 67 Million In November*. 22 de dezembro de 2011 [Online] Disponível em: http://techcrunch.com/2011/12/22/googlesplus/?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+Techcrunch+%28TechCrunch%29 [Acedido em 2 de maio de 2012]
- [33] ComScore – *People Spent 6.7 Billion Hours on Social Networks in October*. 4 de janeiro de 2012 [Online] Disponível em: <http://www.comscoredatamine.com/2012/01/people-spent-6-7-billion-hours-on-social-networks-in-october/>

- [34] Crupi, J., Warner, C. 2009. *Enterprise Mashups: The New Face of Your SOA*. Artigo da *SOA World Magazine* [Online] Disponível em: <http://soa.sys-con.com/node/719917> [Acedido em 3 de julho de 2012]
- [35] Williams, B. 2008. *Enterprise Data Mashups Data Model* [Online] Disponível em: http://www.databaseanswers.org/data_models/enterprise_data_mashups/index.htm [Acedido em 2 de julho de 2012]
- [36] The Open Group, 1995-2012. *SOA Reference Architecture Technical Standard* [Online] Disponível em: http://www.opengroup.org/soa/source-book/soa_refarch/ [Acedido em 2 de janeiro de 2012]
- [37] Kuba, M., Krajíček, O. *Literature search on SOA, Web Services, OGSA and WSRF*. 17 de abril de 2007
- [38] Junic, M., *A Hands-on Introduction to BPEL* [Online] Disponível em <http://www.oracle.com/technetwork/articles/matjaz-bpel1-090575.html> [Acedido em 6 de abril de 2012]
- [39] Chiu, D., Hall, T., Kabir, F., Agrawal, G. *An Approach towards Automatic Workflow Composition through Information Retrieval*. Em atas do 15^o *International Database Engineering and Applications Symposium (IDEAS'11)*. 2011. Lisboa, Portugal.
- [40] Tendeiro, P., Filipe, P., *Practical Information on Sustainable Transport*. *Conference on Electronics Telecommunications and Computers (CETC)*, novembro de 2011. Lisboa, Portugal.
- [41] OASIS Standard (2007). *Web Services Business Process Execution Language Version 2.0* [Online] Disponível em: <http://docs.oasis-open.org/wsbpel/2.0/wsbpel-v2.0.pdf> [Acedido em 3 de abril de 2012]
- [42] Wikimedia Foundation Inc., 2012. *Wikipedia: The Free Encyclopedia* [Online] Disponível em: <http://www.wikipedia.org> [Acedido em 4 de abril de 2012]